

生命科学实验指南系列

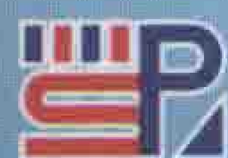


DNA Microarrays: a molecular cloning manual

DNA 微阵列实验指南

〔美〕 D. 鲍特尔 J. 萨姆布鲁克 等 著

吕 华 陆祖宏 孙 啸 译



科 学 出 版 社

(Q-1739.31)



生命科学实验指南系列·典藏版

图解微生物实验指南
免疫学技术及其应用
生物衰老：研究方法与实验方案
精编细胞生物学实验指南
植物蛋白质组学实验指南
蛋白质纯化指南（原书第二版）
环境基因组学实验指南
实验动物血液生理生化参考手册
生理学实验指南
精编免疫学实验指南
酵母遗传学方法实验指南
人干细胞培养
抗体制备及使用实验指南
病毒的电子显微学研究
植物生物学与生态学实验
神经生物学实验原理与技术
DNA微阵列实验指南
基因转移：DNA和RNA的转运与表达
生物实验室管理手册（原书第二版）
精编人类遗传学实验指南
单分子技术实验指南
现代蛋白质工程实验指南
活细胞成像（原书第二版）
遗传变异分析实验指南
表皮细胞实验指南
分子克隆实验指南（原书第三版）（上下册）
精编分子生物学实验指南（原书第五版）
现代神经科学研究技术
生命科学实验设计指南
现代生物化学与分子生物学仪器与设备
分子细胞遗传学——技术和应用
精编蛋白质科学实验指南
实验细胞资源的描述标准与管理规范
实验动物设施运行管理指南
元基因组学：方法和步骤（影印版）
现代工业微生物学实验技术
真核生物转录调控——概念策略与技术（原书第二版）
动物细胞培养——基本技术和特殊应用指南（原书第六版）



科学出版中心 生物分社
联系电话：010-64012501
E-mail：lifescience@mail.sciencep.com
网址：http://www.lifescience.com.cn

销售分类建议：分子生物技术



赛拉艾芙
生命科学订阅号



本书彩图请扫码

www.sciencep.com



ISBN 978-7-03-047486-5

定价（全套）：4500.00元

生命科学实验指南系列·典藏版

DNA 微阵列实验指南

DNA Microarrays：a molecular cloning manual

〔美〕D. 鲍特尔
J. 萨姆布鲁克 等 著
吕 华 陆祖宏 孙 啸 译

科学出版社
北 京

图字：01-2003-4114 号

内 容 简 介

"生命科学实验指南系列"图书均出自名家，包括众多从 Cold Spring Harbor Laboratory Press 和 John Wiley & Sons 等国际知名出版社引进的实验室必备工具书，是生命科学领域最先进、实用、权威的实验手册类优秀图书。该系列图书简单明了，囊括了全世界最著名的生物类实验室操作方法，无论是初学者还是需要深入研究的科研工作者都能从中获益。该系列图书在读者群中有较高的知名度和美誉度，特别是以《分子克隆实验指南》和《精编分子生物学实验指南》为代表，堪称经典，分别被喻为生命科学领域的"蓝宝书"和"红宝书"。现挑选其中的精品集结成典藏版。

图书在版编目（CIP）数据

生命科学实验指南系列：典藏版/雷东锋等编著. —北京：科学出版社，2016

ISBN 978-7-03-047486-5

Ⅰ.①生… Ⅱ.①雷… Ⅲ. ①生命科学–实验–指南 Ⅳ.①Q1-0

中国版本图书馆 CIP 数据核字(2016)第 043878 号

责任编辑：王 静 李 悦

责任印制：张 伟 / 封面设计：刘新新

科学出版社 出版

北京东黄城根北街 16 号

邮政编码：100717

http://www.sciencep.com

北京厚诚则铭印刷科技有限公司 印刷

科学出版社发行 各地新华书店经销

*

2016 年 7 月第 一 版 开本：787×1092 1/16

2016 年 7 月第一次印刷 印张：1310 1/2

字数：31 074 000

定价：4500.00 元

(如有印装质量问题，我社负责调换)

译 者 序

DNA 微阵列的最初构想来自于开发一种高通量的分析工具，可以使用杂交测序（SBH，sequencing by hybridization）的方法对人类基因组进行序列测定。20 世纪 90 年代中期，美国斯坦福大学的 M. Schena 博士应用机器自动点样方法在玻片上固定 DNA 探针，并通过双色荧光标记检测技术同时测定了多个基因的表达。1996 年，美国 Affymatrix 公司应用微电子制造工业领域的照相平版印刷工艺，制备了第一块基因芯片（Genechip®）。在此后 10 多年的发展历程当中，DNA 微阵列的制作工艺日臻成熟，应用领域也在不断拓展，并派生出了蛋白质芯片、细胞和组织芯片等一系列微阵列，以及液相芯片和微流体芯片等新型的芯片分析形式。近几年来，生物芯片技术在新一代高通量 DNA 测序方法中也得到了广泛的研究和应用。总之，在当前的生命科学研究中，微阵列技术提供了一种以往任何技术都无法比拟的高通量、低成本和快速的分析手段。随着人类基因组计划的完成和后基因组时代的到来，生命科学研究的模式面临着重大的变革，从定性到定量研究，从单个因素分析到系统性、全局性的分析，基于微阵列的各种分析方法必将在其中扮演着重要的角色，极大地提高我们获取各种生物信息的能力。

同时我们欣喜地看到，微阵列技术已经走出实验室，走向市场。微阵列和相关设施的商品化，使得更多的人可以接触和使用到这一新兴技术，更多的科研和应用项目也必将从中受益，一个巨大的微阵列产业正在形成。

冷泉港实验室出版社出版的《DNA 微阵列实验指南》一书汇集了这一领域众多的知名专家和学者，他们根据各自研究成果，认真地给出了一系列全面而翔实的 DNA 微阵列实验方案。书中每个方案不仅含有实验的背景知识以及具体的操作内容，此外还提供了大量的信息栏和疑难问题解答，以帮助读者理解实验原理和进行正确的实验操作。全书内容新颖、图文并茂，兼顾了基础理论和应用实验技术，因而不仅仅是一本有价值的实验手册，同时也是了解和学习 DNA 微阵列技术的不可多得的、权威性的参考书。科学出版社将该书介绍给国内读者，对于推动我国生物芯片的研究与应用将会具有非常积极的作用。

感谢东南大学生物电子学实验室白云飞、程璐、董献军、高璐、葛芹玉、顾万君、郭晓英、贺全国、侯鹏、祭美菊、贾可、李同祥、刘和平、刘全俊、刘正春、梅茜、王宏、王进科、王志飞、肖鹏峰、肖忠党、谢建明、谢雪英、杨锡南、杨玉志、张春秀和周士新等参与了本书的最初翻译工作。为保持全书译文的准确性和前后内容的一致性以及连贯性，由吕华、陆祖宏和孙啸三位老师经大量修改润色，最终定稿。

科学出版社的李悦、王静老师为本书的出版工作付出了大量努力，在此谨向她们表示由衷的感谢！

最后感谢我们的同事和家人在本书翻译工作期间给予的理解、鼓励和支持。

希望本书的出版会对 DNA 微阵列领域的科研和应用人员有所帮助。由于译者能力所限，翻译中的不妥之处，恳请读者批评指正。

译　者

2007 年 12 月于东南大学

原 著 序

基因组学（genomics）出现于20世纪70年代，当时Sanger提出一个观点，认为确定个体的全部基因组序列是理解其生物学的良好起点。他与剑桥分子生物学实验室的同事一起，用Sanger测序法[1]测定了ϕX174噬菌体[2]、人线粒体[3]和λ噬菌体[4]的基因组序列。根据其他一些发现，ϕX174序列揭示了重叠基因，线粒体基因则表明其中使用了交替密码子。这个开创性的工作推动了更大的项目，最终促成近年来对人类基因组的大部分序列进行测定[5,6]*。公共数据库中数量巨大的高质量序列信息，作为对Sanger遗产的衡量，也向我们提出了一个挑战，即从序列到基因组功能和高度复杂的生物体。基于计算机的分析方法也由Sanger研究组开创[7]。例如，通过发现候选基因，从序列出发经过很长过程来提取生物学上相关的信息。但是计算机无法测量基因表达的水平，或是分析DNA样品中序列的改变。对于这些和其他问题，我们不仅需要改进现有的实验方法，同时也需要开创在规模上能与复杂基因组的巨大尺寸相适应的方法。

有两种方法可以大规模地进行我们的实验：自动串行处理是将Sanger DNA测序法改进以适合分析复杂基因组需要的巨大规模而采用的方法；另一种方法是并行处理，以微阵列为例，它具有很多优点，但它需要有一种简单过程，以便于发展成为携带许多分立测试部位的装置。核酸从单独两条碱基序列互补的链形成双链，正是这样一种过程。包含大量已知序列探针组的微阵列，可以使得许多杂交反应在一个小型装置中并行地进行。

微阵列的基本设计有许多变化，但几乎全部使用非渗透性的支持物作为基底，其上固定探针样点。与基于分子杂交早期分析方法中使用的膜相比，非渗透性支持物有许多优点。重要的是，它使得在一个非渗透性的表面聚集许多细微部件成为可能。典型的微阵列结构的单元尺寸约为20～200 μm，因此数千个结构单元可以安放在标准的显微镜载玻片上。不仅如此，由于分析物溶液不能透过表面，因此杂交和漂洗过程要比可渗透支持物快得多。玻璃具有所希望的物理特性，因此是最常用的支持物[8]。与通常的印迹膜不同，玻璃在润湿和干燥时不会变形。阵列的制造和读出过程中，空间稳定性对于探针点的精确定位极为重要。

适合于阵列制备的印刷方法能够分辨到微米级，并可用于传送预先合成的探针、cDNA克隆、PCR产物或合成的寡核苷酸，或者是传送试剂用于原位合成探针。对大多数自行制备阵列的研究人员来说，用预先合成的探针点样是唯一可使用的方法。原位方法有几家厂商在使用，但这种方法目前花费昂贵，大多数用户无法使用，尤其是那些在大学实验室里工作的人。

---

*2003年4月14日，美国联邦国家人类基因组研究项目负责人弗朗西斯·柯林斯博士在华盛顿宣布，美、英、日、法、德和中国科学家经过13年努力共同绘制完成了人类基因组序列图，人类基因组计划所有目标全部实现。——译者注

某些应用，如分析来自酿酒酵母（*Saccharomyces cerevisiae*）的基因表达，需要一套标准探针——涵盖全部基因产物的探针。一旦建立起一个最优探针集，采用它作为标准有许多优点。用预先合成的探针点样[9]和照相平版印刷方法[10]适合于生产许多在设计上相同的阵列；然而它们不适合一轮又一轮生产不同的阵列，这在一些研究计划中是必需的。另一方面，原位喷印合成[11]和应用微镜阵列的光导制备[12]不仅适合于重复生产一种设计的阵列，而且也适合于生产具有不同设计的阵列，只要简单改变从计算机到喷墨打印机或微镜阵列的指令集就能实现。

对所有制备方法而言，质量控制是一个主要的问题。很少有方法可以检查每个点上的序列及放置或合成探针的数量[13, 14]。通过反复扩增增加探针数量可能会导致杂质序列的过度增加。由于该原因，化学合成探针可能会取代克隆和 PCR 产物。对选择性剪接作为调节基因功能机制重要性的逐步认识也显示，必须使用寡核苷酸按外显子逐个测量基因表达。原位合成产生的探针质量主要取决于耦联化学的效率。光合成方法每步的产率相对要低[15]，并限制了寡核苷酸的长度只能到 25 聚体。但是喷墨制备可以产生长度为 60 聚体的寡核苷酸[11]，这对大多数应用是足够的。样点质量是影响阵列性能的另一个因素。例如，点样针或者毛细管可能会产生形状和尺寸差异极大的点，使得寻找和积分杂交信号工作更加困难。

所有方法的应用都希望达到高灵敏度和特异性的效果，但这可能导致互相矛盾。短探针，大约 20 聚体，杂交信号相对低；较长的探针灵敏度较高，但可能特异性较低，因为它们可能会与序列相似度较低的靶分子相互作用。一项详细研究[11]表明杂交条件对特异性和灵敏度的巨大影响，并强调了这两种需求之间的矛盾——最佳特异性需要的条件比给出最佳杂交结果条件的严格度更高。在一些应用中，必须使用短探针。例如，通过杂交检测单碱基错配，不可能使用长度大于约 20 个碱基的探针，较长的探针几乎显示不出完全匹配双链和错配双链在结果上的差异。

微阵列实验的巨大规模需要计算机的帮助。在实验开始时，设计阵列和引导制备过程需要计算机；计算机用于获取杂交的靶分子图像；将图像转换成对杂交程度的有效度量需要计算机；将杂交程度解释为对靶分子中互补序列数量有意义的测量需要计算机。已有软件包可用于处理该过程的所有方面。要求最高的计算任务是结果的解释，大量精力花在用统计学评价结果的质量和发现数据中模式的方法上，如应用聚类算法去发现基因调控的公共模式。一些商用软件包对大学用户来说过于昂贵，但类似从 EBI http://www.ebi.ac.uk/microarray 获取的公共域程序可以解决这一需求。

微阵列技术的成功推动了它的改进，如更加灵敏地检测 mRNA。反过来，这种要求也导致了需要对内在过程的机制进行更加深入的了解。作为回应，我们开始看到发表的研究工作越来越多是关于固定的探针结构和杂交机制，这对器件的性能十分重要。例如，原位合成可以形成非常高的寡核苷酸密度，以致表面上的分子之间会相互作用；空间位阻抑制了靶分子和探针的紧密靠近。空间效应可以通过长的手臂分子固定探针[13,16]，或是减小它们的密度[13]来得到缓解，并可使杂交效果增加 6～150 倍。高的探针密度对杂交动力学也有影响。典型的探针密度约为 1～10 pmol/mm^2。这样一个直径为 100 μm 的典型样点中有 0.01～0.1 pmol（约 10^{10}～10^{11} 个分子）探针。每个阵列点中探针的量通常远远超出溶液中它的靶分子数量；在一个典型的转录水平分析中，总共

约 1 μg RNA（约 1 pmol）可能与含有10 000个点的阵列杂交，其中共有约 1 nmol 探针。这直接导致位于样点上方的靶分子溶液中与样点探针互补的靶分子迅速耗尽，反应是受扩散控制的，搅拌对整个速度有巨大的影响[17]。一个实用的推论是将靶分子稀释至可以有效搅拌的大容量后再加到阵列上，这比使用不能搅拌的浓缩薄层要好。

微阵列技术已经在不相同的应用领域取得了许多重要的结果。转录物分析被用于发现外显子和基因，从而对人类基因组序列草图进行注释[18]。对基因组 DNA 分析，检测到肿瘤中的扩增和缺失[19, 20]。差异基因表达分析揭示了公共调节途径中的基因网络[21, 22]，它显示了肿瘤之间不能为常规方法所区分的差异[23, 24]。这些重要结果没有一项可以用其他分析方法如此简单或迅速地得到。

阵列平台是一个相对复杂的技术，它将来自许多学科、充满活力、有共同兴趣的人集中在一起：除了分子生物学家、遗传学家和计算机科学家以外，还有工程师、材料科学家、数学家和化学家。未来将有新的方法和用途涌现出来。已经在使用这项技术的人要跟上发展的步伐；而该领域的新进入者则需要得到帮助。这本书提供的最新信息，将会引导用户及那些寻求改进该方法和开发新用途的人员继续努力实践。

Ed Southern

序言参考文献

1. Sanger F. and Coulson A.R. 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J. Mol. Biol.* **94:** 441–448.
2. Sanger F., Coulson A.R., Friedmann T., Air G.M., Barrell B.G., Brown N.L., Fiddes J.C., Hutchison C.A., III, Slocombe P.M., and Smith M. 1978. The nucleotide sequence of bacteriophage phiX174. *J. Mol. Biol.* **125:** 225–246.
3. Anderson S., Bankier A.T., Barrell B.G., de Bruijn M.H., Coulson A.R., Drouin J., Eperon I.C., Nierlich D.P., Roe B.A., Sanger F., Schreier P.H., Smith A.J., Staden R., and Young I.G. 1981. Sequence and organization of the human mitochondrial genome. *Nature* **290:** 457–465.
4. Sanger F., Coulson A.R., Hong G.F., Hill D.F., and Petersen G.B. 1982. Nucleotide sequence of bacteriophage lambda DNA. *J. Mol. Biol.* **162:** 729–773.
5. Lander E.S., Linton L.M., Birren B., Nusbaum C., Zody M.C., Baldwin J., Devon K., Dewar K., Doyle M., FitzHugh W., et al. 2001. Initial sequencing and analysis of the human genome. *Nature* **409:** 860–921.
6. Venter J.C., Adams M.D., Myers E.W., Li P.W., Mural R.J., Sutton G.G., Smith H.O., Yandell M., Evans C.A., Holt R.A., et al. 2001. The sequence of the human genome. *Science* **291:** 1304–1351.
7. Staden R. 1977. Sequence data handling by computer. *Nucleic Acids Res.* **4:** 4037–4051.
8. Maskos U. and Southern E.M. 1992. Oligonucleotide hybridizations on glass supports: A novel linker for oligonucleotide synthesis and hybridization properties of oligonucleotides synthesised in situ. *Nucleic Acids Res.* **20:** 1679–1684.
9. Schena M., Shalon D., Davis R.W., and Brown P.O. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270:** 467–470.
10. Lipshutz R.J., Fodor S.P., Gingeras T.R., and Lockhart D.J. 199. High density synthetic oligonucleotide arrays. *Nat. Genet.* (suppl.) **21:** 20–24.
11. Hughes T.R., Mao M., Jones A.R., Burchard J., Marton M.J., Shannon K.W., Lefkowtiz S.M., Ziman M., Schelter J.M., Meyer M.R., Kobayashi S., Davis C., Dai H., He Y.D., Stephaniants S.B., Cavet G., Walker W.L., West A., Coffey E., Shoemaker D.D., Stoughton R., Blanchard A.P., Friend S.H., and Linsley P.S. 2001. Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. *Nat. Biotechnol.* **19:** 342–347.
12. Singh-Gasson S., Green R.D., Yue Y., Nelson C., Blattner F., Sussman M.R., and Cerrina F. 1999. Maskless fabrication of light-directed oligonucleotide microarrays using a digital micromirror array. *Nat. Biotechnol.* **17:** 974–978.

13. Peterson A.W., Heaton R.J., and Georgiadis R.M. 2001. The effect of surface probe density on DNA hybridization. *Nucleic Acids Res.* **29:** 5163–5168.
14. Gray D.E., Case-Green S.C., Fell T.S., Dobson P.J., and Southern E.M. 1997. Ellipsometric and interferometric characterization of DNA probes immobilized on a combinatorial array. *Langmuir* **13:** 2833–2842.
15. McGall G.H., Barone A.D., Diggelmann M., Fodor S.P.A., Gentalen E., and Ngo N. 1997. The efficiency of light directed synthesis of DNA arrays on glass substrates. *J. Am. Chem. Soc.* **119:** 5081–5090.
16. Shchepinov M.S., Case-Green S.C., and Southern E.M. 1997. Steric factors influencing hybridisation of nucleic acids to oligonucleotide arrays. *Nucleic Acids Res.* **25:** 1155–1161.
17. Ramakrishnan R., Dorris D., Lublinsky A., Nguyen A., Domanus M., Prokhorova A., Gieser L., Touma E., Lockner R., Tata M., Zhu X., Patterson M., Shippy R., Sendera T.J., and Mazumder A. 2002. An assessment of Motorola CodeLink microarray performance for gene expression profiling applications. *Nucleic Acids Res.* **30:** e30.
18. Shoemaker D.D., Schadt E.E., Armour C.D., He Y.D., Garrett-Engele P., McDonagh P.D., Loerch P.M., Leonardson A., Lum P.Y., Cavet G., Wu L.F., Altschuler S.J., Edwards S., King J., Tsang J.S., Schimmack G., Schelter J.M., Koch J., Ziman M., Marton M.J., Li B., Cundiff P., Ward T., Castle J., Krolewski M., Meyer M.R., Mao M., Burchard J., Kidd M.J., Dai H., Phillips J.W., Linsley P.S., Stoughton R., Scherer S., and Boguski M.S. 2001. Experimental annotation of the human genome using microarray technology. *Nature* **409:** 922–927.
19. Fritz B., Schubert F., Wrobel G., Schwaenen C., Wessendorf S., Nessling M., Korz C., Rieker R.J., Montgomery K., Kucherlapati R., Mechtersheimer G., Eils R., Joos S., and Lichter P. 2002. Microarray-based copy number and expression profiling in dedifferentiated and pleomorphic liposarcoma. *Cancer Res.* **62:** 2993–2998.
20. Hodgson G., Hager J.H., Volik S., Hariono S., Wernick M., Moore D., Nowak N., Albertson D.G., Pinkel D., Collins C., Hanahan D., and Gray J.W. 2001. Genome scanning with array CGH delineates regional alterations in mouse islet carcinomas. *Nat. Genet.* **29:** 459–464.
21. Miki R., Kadota K., Bono H., Mizuno Y., Tomaru Y., Carninci P., Itoh M., Shibata K., Kawai J., Konno H., Watanabe S., Sato K., Tokusumi Y., Kikuchi N., Ishii Y., Hamaguchi Y., Nishizuka I., Goto H., Nitanda H., Satomi, S., Yoshiki A., Kusakabe M., DeRisi J.L., Eisen M.B., Iyer V.R., Brown P.O., Muramatsu M., Shimada H., Okazaki Y., and Hayashizaki Y. 2001. Delineating developmental and metabolic pathways in vivo by expression profiling using the RIKEN set of 18,816 full-length enriched mouse cDNA arrays. *Proc. Natl. Acad. Sci.* **98:** 2199–2204.
22. Zhu G., Spellman P.T., Volpe T., Brown P.O., Botstein D., Davis T.N., and Futcher B. 2000. Two yeast forkhead genes regulate the cell cycle and pseudohyphal growth. *Nature* **406:** 90–94.
23. Sorlie T., Perou C.M., Tibshirani R., Aas T., Geisler S., Johnsen H., Hastie T., Eisen M.B., van de Rijn M., Jeffrey S.S., Thorsen T., Quist H., Mmatese J.C., Brown P.O., Bostein D., Eystein Lonning P., and Borresen-Dale A.L. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci.* **98:** 10869–10874.
24. Alizadeh A.A., Eisen M.B., Davis R.E., Ma C., Lossos I.S., Rosenwald A., Boldrick J.C., Sabet H., Tran T., Yu X., Powell J.I., Yang L., Marti G.E., Moore T., Hudson J., Jr., Lu L., Lewis D.B., Tibshirani R., Sherlock G., Chan W.C., Greiner T.C., Weisenburger D.D., Armitage J.O., Warnke R., Staudt L.M., et al. 2000. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* **403:** 503–511.

前　　言

当微阵列界就如何产生可靠结果所需的实验类型和分析方法这一问题趋于共识的时候，本书开始酝酿——这一情形令人回想起重组 DNA 研究的早期阶段。与早期克隆实验一样，点样微阵列技术的早期使用者发现很少有完美的实验方案，并且几乎没有什么商品化的试剂。在这个开拓时期，技术的发展大多数是由在个别大学实验室中工作、使用定制仪器和“自制”试剂的博士后和学生所推动的。我们特别幸运，能有许多这一年轻领域的领头人参与本书的写作。

我们的目标是出版一本实验指南，使得任何持之以恒和细心的人，若有适当的预算开支，就可以制造、使用和分析微阵列。在 2001 年 4 月，本书材料的汇编工作逐渐落实，当时 Jan Witkowski 慷慨地答应在冷泉港实验室 Banbury 中心主办一次微阵列会议。参会人员提供并分析了许多实验方案，现已成为本书的核心。在第二年里，最初的草案被重新优化和编辑，这项工作经常是针对这一领域的发展进行实质性的修改，并补充了关键实验方案，以及数量不断增加的微阵列设备和试剂供应商的信息。

用于平行分析基因表达的现代 DNA 微阵列，是生长在薄膜过滤膜上、细菌 cDNA 克隆网格状阵列的直接派生物。由于网格滤膜阵列仍是进行表达分析有效且经济的方法，因而书中的一些实验方案叙述了它的合成与使用。另一方面，本书大部分内容阐释了当前的两种主要技术：由斯坦福大学 Patrick Brown 实验室首创的在玻璃载玻片上的点样 DNA 微阵列和由 Affymetrix 公司制造的原位合成寡核苷酸阵列。尽管需要一定的努力和技巧，但点样微阵列是可以在大学实验室中制造的。与此相比，Affymetrix 阵列的使用更加接近“现成型”，它回避了制造阵列本身，并且使用试剂盒进行靶分子标记和杂交。因为这两种技术都有一些难点，本书通过详细的“疑难解答”部分，来讨论点样和 Affymetrix 阵列中遇到的常见问题。经慎重考虑，本书着重于使用 cDNA 阵列进行表达分析，因为这是迄今为止该技术的主要应用。为了便于将此分析方法用于极少量纯化细胞，本书介绍关于微分离和少量 RNA 体外扩增的实验方案。一旦技术得以发展，人类的智慧同样会赋予它新的用途，这当然也是微阵列领域的一个特点。使用微阵列检测基因组的变化——自然的和病源的——日益增多，本书用较多篇幅介绍了这类应用。微阵列产生大量必须由其他方法验证的数据，因此本书还包含了应用原位杂交和免疫组化对微阵列数据进行快速下游分析的实验方案。

发展有力的和用户友好的生物信息学工具来管理微阵列实验，挖掘与提取数据仍然是一个主要前沿。这些工具正处在迅速发展的时期。应用集中在非监督的数据聚类方法，如 Mike Fisen 的聚类程序和相似的程序，对于探索几乎没有其他信息可用于引导分析的微阵列数据来说，仍然非常有效。然而，基于有监督机器学习的方法正逐渐变得有力，并为未来发展提供了广阔的领域。倘若微阵列生物信息学发展迅速，那么没有一种关于该领域的描述会长期有效或者永不过时。本书中详细介绍了已经建立的技术，包括层次聚类法。然而，这部分的作者将大多数内容侧重于不易改变的基本概念，如获得

数值数据、实验设计和数据库的要求。

由于微阵列领域变化极为迅速，因此，将初步实验方案迅速出版成书很重要。我们向参编者表示感谢，他们非常慷慨地接受了紧迫的时间限制，并且很专业地完成了编写工作。这当中许多提供单独实验方案的科学家出现在拍摄于 Banbury 会议的照片中*。他们和其他许多人在本书编辑的不同时期提供了无法估价的、持续的技术建议。我们特别感谢 Vivek Mittal，Marcelo Bento Soares，Christoph Ileller，Roland Toiler，Josh Dabna 和 Karin Schutze。Ed Southern 在某种意义上来说是这一领域的创始者，感谢他为本指南撰写了优美的序。

在本书制作中，我们得到了各种各样的支持。我们感谢 Beth Nickerson，他在本指南接近后期制作时加入，提供了许多无价的洞察性建议。尽管远在半个世界之遥，冷泉港实验室出版社的编辑和制作员工令人难忘的高效工作和投入，将出版过程的每个步骤变得极为轻松。我们尤其要感谢 David Crotty，Dorothy Brown，Susan Schaefer，Nora Rice，Mary Cozza，Sian Curtis，Denise Weiss 和 Kimberly LaVine，他们的成果得到了 Inez Silaliano 和 Jan Argentine 的协调和强化，在似乎有无数的工作有待完成时，他们两人为大家提供了智慧、镇定的指导。我们感谢在 Peter MacCallum 癌症研究所的同事，当会议取消时他们给予宽容和理解，当无数次被告知“本书即将完成”时他们始终表现出坚信不疑。Linda Stevens 给予了出色的帮助，他负责重新安排会议和按常规管理我们的活动。

特别感谢 Kaaren Janssen，离开她就不可能有这本书。从最初组织 Banbury 会议，到编辑最后一个实验方案，她是最活跃的参加者之一，在整个过程中尽力保持欢乐、热情、不知疲倦和毫无怨言。我们特别幸运能和她一起工作，并享有她的友谊。最后，我们向我们的家庭表示衷心的感谢，在许多失去的夜晚和缺席的周末，他们给予了我们支持、鼓励和理解。

D. Bowtell
J. Sambrook

* Banbury 会议照片见封底。

原著作者名单

Donna Albertson *Comprehensive Cancer Center, Cancer Research Institute, and Department of Laboratory Medicine, University of California, San Francisco*

Catherine A. Ball *Stanford University School of Medicine, Stanford, California*

Bernd Becker *Universität Regensburg, Regensburg, Germany*

Kevin G. Becker *DNA Array Unit, National Institute on Aging, National Institutes of Health, Baltimore, Maryland*

Monique Bernsen *Academisch Ziekenhuis Nijmegen, Nijmegen, The Netherlands*

Carolyn J.M. Best *Pathogenetics Unit, Laboratory of Pathology and Urologic Oncology Branch, National Cancer Institute, Bethesda, Maryland*

Tone Björnsen *University Hospital, Department of Pathology, Regionsykehuset Tromso-RiTo, Tromsoe, Norway*

Michael Bittner *National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland*

Karl Brand *Peter MacCallum Cancer Institute, East Melbourne, Victoria, Australia*

Dieter Brocksch *Business Group Microscopy, Carl Zeiss, Göttingen, Germany*

Renate Burgemeister *P.A.L.M. Microlaser Technologies AG, Bernried, Germany*

Christer Bush *University Hospital, Department of Pathology, Regionsykehuset Tromso-RiTo, Tromsoe, Norway*

Ron Butow *University of Texas Southwestern Medical Center, Dallas*

Piero Carninci *Genome Science Laboratory, Riken Main Campus, Wako, Saitama, Japan*

Settara Chandrasekharappa *National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland*

Chris Cheadle *DNA Array Unit, National Institute on Aging, National Institutes of Health, Baltimore, Maryland*

Yidong Chen *National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland*

Geoffrey Childs *Albert Einstein College of Medicine, Bronx, New York*

Rodrigo F. Chuaqui *Pathogenetics Unit, Laboratory of Pathology and Urologic Oncology Branch, National Cancer Institute, Bethesda, Maryland*

Annette Clement-Sengewald *Frauenklinik der Ludwig Maximilians Universität, Munich, Germany*

Suzanne Dee *Affymetrix, Inc., Santa Clara, California*

Joseph DeRisi *University of California, San Francisco*

Maximilian Diehn *Stanford University School of Medicine, Stanford, California*

James Eberwine *University of Pennsylvania Medical Center, Philadelphia*

Michael R. Emmert-Buck *Pathogenetics Unit, Laboratory of Pathology and Urologic Oncology Branch, National Cancer Institute, Bethesda, Maryland*

Charles B. Epstein *Aventis Pharmaceuticals, Cambridge, Massachusetts*

Jian-Bing Fan *Affymetrix, Inc., Santa Clara, California*

Gabriele Friedemann *P.A.L.M. Microlaser Technologies AG, Bernried, Germany*

John W. Gillespie *Pathogenetics Unit, Laboratory of Pathology and Urologic Oncology Branch, National Cancer Institute, Bethesda, Maryland*

Richard Glynne *Eos Biotechnology, South San Francisco, California*

Sean Grimmond *ARC-SRC for Functional and Applied Genomics, Gene Expression Profiling Facility, Institute of Molecular Bioscience, University of Queensland St Lucia, Australia*

Thomas Harris *Albert Einstein College of Medicine, Bronx, New York*

Wolfgang Heckl *Institute of Cristallography, Ludwig Maximilians Universität, Munich, Germany*

Joel N. Hirschhorn *Whitehead Institute/MIT Center for Genome Research, Cambridge, Massachusetts; Harvard Medical School, Boston, Massachusetts; Divisions of Genetics and Endocrinology, Children's Hospital, Boston, Massachusetts*

Andrew Holloway *Peter MacCallum Cancer Institute, East Melbourne, Australia*

Galen Hostetter *Cancer Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland*

Bi-Huei Hou *Carnegie Institute of Washington, Stanford, California*

Xiaohua Huang *Affymetrix, Inc., Santa Clara, California*

Trey Ideker *Whitehead Institute for Biomedical Research, Cambridge, Massachusetts*

Vishwanath Iyer *University of Texas, Austin*

Janet Estee Kacharmina *University of Pennsylvania Medical Center, Philadelphia*

Olli-P. Kallioniemi *Cancer Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland*

Sophie Katsabanis *Peter MacCallum Cancer Research Institute, Melbourne, Australia*

Javed Khan *Advanced Technology Center, Pediatric Oncology Branch, National Cancer Institute, NIH, Gaithersburg, Maryland*

Juha Kononen *Cancer Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland*

Georgia Lahr *Academic-Hospital München-Harlaching, Munich, Germany*

Jerry Lee *Eos Biotechnology, South San Francisco, California*

Isabel M. Leiva *Pathogenetics Unit, Laboratory of Pathology and Urologic Oncology Branch, National Cancer Institute, Bethesda, Maryland*

Yin-Mei Lim *Whitehead Institute/MIT Center for Genome Research, Cambridge, Massachusetts*

Per Lindahl *Göteborg University, Department of Medical Biochemistry, Göteborg, Sweden*

Kerstin Lindblad-Toh *Whitehead Institute/MIT Center for Genome Research, Cambridge, Massachusetts*

Chih Long Liu *Biological and Biomedical Sciences, Division of Biomedical Sciences, Harvard Medical School, Boston, Massachusetts*

Robert Lucito *Cold Spring Harbor Laboratory, Cold Spring Harbor, New York*

Tracy McIntosh *University of Pennsylvania Medical Center, Philadelphia*

Paolo Marciano *University of Pennsylvania Medical Center, Philadelphia*

Francesco M. Marincola *HLA and Immunogenetics Laboratory, National Cancer Institute, National Institutes of Health, Bethesda, Maryland*

Aldo Massimi *Albert Einstein College of Medicine, Bronx, New York*

Hajime Matsuzaki *Affymetrix, Inc., Santa Clara, California*

Annette Mayer *Institute of Anthropology and Human Genetics, Goethe University, Frankfurt/Main, Germany*

Damares Monte *Carnegie Institute of Washington, Stanford, California*
Maria Murphy *Peter MacCallum Cancer Institute, East Melbourne, Australia*
Sven Nilsson *Göteborg University, Department of Medical Biochemistry, Göteborg, Sweden*
Norma J. Nowak *School of Medicine and Biomedical Sciences, State University of New York, Buffalo and Roswell Park Cancer Institute, Buffalo, New York*
Sujatha Panavally *National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland*
Nila Patil *Affymetrix, Inc., Santa Clara, California*
Daniel Pinkel *Comprehensive Cancer Center and Department of Laboratory Medicine, University of California, San Francisco*
Jonathan R. Pollack *Stanford University School of Medicine, Stanford, California*
Virgil Rhodius *University of California, San Francisco*
Thomas Ryder *Affymetrix, Inc., Santa Clara, California*
Guido Sauter *Institute of Pathology, University of Basel, Switzerland*
Stefan J. Scheidl *Göteborg University, Department of Medical Biochemistry, Göteborg, Sweden*
Karin Schütze *P.A.L.M. Microlaser Technologies AG, Bernried, Germany*
Richard Segraves *Comprehensive Cancer Center, Cancer Research Institute, University of California, San Francisco*
Gavin Sherlock *Stanford University School of Medicine, Stanford, California*
Pamela Sklar *Whitehead Institute/MIT Center for Genome Research, Cambridge, Massachusetts; Harvard Medical School, Boston, Massachusetts; Department of Psychiatry, Massachusetts General Hospital, Boston*
Antoine Snijders *Comprehensive Cancer Center, Cancer Research Institute, University of California, San Francisco*
Shauna Somerville *Carnegie Institute of Washington, Stanford, California*
Terry Speed *University of California, Berkeley*
Paul T. Spellman *University of California, Berkeley*
Monika Stich *Academic-Hospital München-Harlaching, Munich, Germany*
Wilhelm Stolz *Universität Regensburg, Regensburg, Germany*
Minoru Takemoto *Göteborg University, Department of Medical Biochemistry, Göteborg, Sweden*
Stefan Thalhammer *Institute of Cristallography, Ludwig Maximilians Universität, Munich, Germany*
Ed Tom *Eos Biotechnology, South San Francisco, California*
Melanie Trivett *Peter MacCallum Cancer Institute, East Melbourne, Australia*
Marcory C.R.F. van Dijk *Academisch Ziekenhuis Nijmegen, Nijmegen, The Netherlands*
Thomas Vogt *Universität Regensburg, Regensburg, Germany*
Ena Wang *HLA and Immunogenetics Laboratory, National Cancer Institute, National Institutes of Health, Bethesda, Maryland*
Jun S. Wei *Advanced Technology Center, Pediatric Oncology Branch, National Cancer Institute, NIH, Gaithersburg, Maryland*
Michael Wigler *Cold Spring Harbor Laboratory, Cold Spring Harbor, New York*
Dorian Willhite *Eos Biotechnology, South San Francisco, California*
William H. Wood III *DNA Array Unit, National Institute on Aging, National Institutes of Health, Baltimore, Maryland*
Yee Hwa Yang *University of California, Berkeley*
Suzanne Ybarra *Eos Biotechnology, South San Francisco, California*

目　录

第 1 章　点样微阵列探针的生成

1

导言

2

探针和靶

制备点样 DNA 微阵列的第一步是生成“现成的阵列材料（array-ready material）”，以此作为点样的给料。对用于制备微阵列材料和与微阵列进行杂交的材料的命名，有一些争论不休的观点。本指南中，我们依照惯例，将微阵列上的材料称为探针（probe），

将与微阵列杂交的材料称为靶（target）。该命名的背景和逻辑性将在本导言结尾部分关于**探针和靶的命名**信息栏中叙述。

本节内容着重于介绍制备探针用于基因表达分析，这是到目前为止，DNA 微阵列最具优势的用途。手册的后面几节讲述了微阵列的点样和其他用途探针的制备，如基于微阵列的比较基因组杂交的基因组克隆。

微阵列使用的探针可以通过合成的寡核苷酸或 cDNA 片段得到。对于多数研究者来说，用于表达分析的理想微阵列是一个探针序列经过验证的复杂阵列，其中的每条序列都是唯一的，并且与相关序列的交叉杂交最低。微阵列实验的部分能力在于采用了一种无偏差的基因发现方法（Brown and Botstein 1999）。因此，理想情况下微阵列上探针是包括可变剪接体在内的基因组已表达部分的详细表示。然而，对于复杂的生物体，要达到这个目标，目前在成本、可利用的合适探针以及制备包含作为数千个外显子代表的微阵列在逻辑上的需求等方面仍存在一定困难。探针材料的来源不同，使得制备“完美的”表达阵列受到约束。

克隆集

对各种组织和生物体的 cDNA 文库进行测序的工作开始于 20 世纪 80 年代末和 90 年代初，目的是辨别转录子，鉴定表达的基因，并在几年内绘出整个基因组中已表达序列区域的图谱。尽管这些宏伟的目标还没有全部实现，然而 cDNA 测序计划已经为生成用于微阵列的复杂探针集合提供了所需要的物理资源，如克隆文库。关于**克隆集**的信息栏中叙述了人和一些模式物种克隆集的来源，并解释了其应用的基本原理。关于克隆集供应商的附带表格（表 1-5）中给出了目前使用的一些克隆集的可用性和复杂性信息。

近来，许多研究者使用现成的克隆集的最新版本进行 cDNA 片段的阵列制备，并为克隆集补充了符合他们特殊需要的特有的表达序列标签（EST）。特有 EST 是通过对数据库的仔细扫描来鉴定的，并可以从供应商处购得。如果克隆集要保持存活、无污染且排放整齐，那么处理成千上万的 cDNA 就是一个特殊的挑战。最初的克隆集会受到噬菌体 T1 污染，单孔中存在多个克隆或错误的序列分配等（Halgren et al. 2001）。在关于**基因组和 cDNA 文库中 T1 噬菌体的预防和检测**信息栏中，给出了解决噬菌体 T1 污染克隆集这一问题可供选择的方法。

3

据报道，在维护良好的克隆集中，1%～5%的克隆发生了错误序列分配，也就是说它们不含应有的序列，目前对这一数字尚有争议（Knight 2001）。单个克隆的重新纯化及其序列验证，虽然不能完全消除但可以减少可用克隆集中的误差率。新误差引入的速率与克隆集供应商和使用克隆集的各个实验室的严谨程度有关。方案 1 列出了严格复制和保存克隆集的步骤。然而，由于从克隆复制到最后微阵列点样中的任何一步都能引起误差，所以不能假定微阵列实验中的结果都是完全可靠的，必须通过独立的方法对结果进行确认。

用于阵列的 DNA 一般通过高通量的聚合酶链反应（PCR）由克隆集制备，而不是通过纯化重组子，如质粒来制备。由于克隆集通常使用范围有限的克隆载体，因而一般能够应用通用引物进行 cDNA 插入片段的扩增（表 1-1）。所制备微阵列的质量关键在

于有效、高产和可重复的 PCR 扩增，并且要确保产物不受制备中使用的引物、脱氧核苷酸和酶的影响。方案 2 介绍了 PCR 装置；方案 3、4、5 详述了扩增产物的纯化、定量和质量控制；方案 5 提供了处理和重悬备用的微阵列材料的细节以进行点样和储存。

目前，酵母是唯一的已知全部基因序列的真核生物。方案 6 介绍了通过基因特异的引物集进行 PCR 扩增从酵母制备新克隆集的技术。随着可用基因组序列数量的增加和分析工具的完备，克隆集的序列复杂性将会增加，所代表的表达基因组的比例也会增加。然而，在此之前仍需要制备额外的克隆集来代表每个特殊物种、组织和发育阶段基因组中表达的部分。关于标准化和扣除过程的信息及制备克隆文库以增加含量较低的 mRNA 种类的常用技术，参见附录 2 和 Bonaldo 等（1996）的文献。

微阵列构建中基因组信息的应用

人类基因组和大量模式生物 DNA 序列的完成使得用计算机设计探针进行表达研究成为可能。一种方法是用短的寡核苷酸引物扩增基因组表达区域的相应片段，这一策略最先应用于酵母（Iyer et al. 1999）。本节结束部分关于**酵母探针**的信息栏介绍了制备包含酿酒酵母（*Saccharomyces cerevisiae*）大部分已表达基因的探针。通过这种方法扩增了其他基因组，包括细菌基因组（Wei et al. 2001）和植物基因组（Hegde et al. 2000）。尽管这是一种吸引人的系统方法，但在设计引物用于高效产出 PCR 扩增产物方面仍然存在着挑战。更为直接的方法是在已知和预测的可读框基础上，用含有 50～70 个碱基的寡核苷酸（所谓“长链寡核苷酸”）直接印刷点样阵列。基因表达区域可使用已有的 cDNA 序列信息和序列预测程序来确定。由于寡核苷酸合成的成本从每个碱基数美元降至数美分，且全长的长链寡核苷酸产量得到提高，因此，制备复杂的寡核苷酸点样微阵列逐渐变得容易。像克隆集一样，包含了不同物种中相当比例表达序列的长链寡核苷酸集可通过购买得到。在关于**长链寡核苷酸的点样：替代用于表达分析的 cDNA 微阵列**信息栏中，对长链寡核苷酸设计时的注意事项和各种方法进行了讨论。相应的表格（表 1-6）中提供了商品化寡核苷酸文库的信息。

探针和靶的命名

25 年来，Northern 印迹法已经成为分析 RNA 表达的主要技术，并强烈影响着我们对探针和靶的理解。在 Northern 印迹法中，探针是标记的，通常是放射性的，游离于溶液中，并且是已知的——如克隆 DNA 的单个片段。靶 RNA 是复杂的，固定于固相支持物的全部待测物。在微阵列实验中，已知物和未知物在物理位置上正好相反。全部的已知物，克隆 DNA 的单个片段排列于固相表面，如载玻片；全部的未知物，如复杂的合并 RNA（RNA pool），是在溶液中。然而实验目的并没有改变，即测定 RNA 样品中单个基因表达的水平。

物理位置的颠倒导致微阵列实验中“探针”的含义有些混乱。一些研究人员用“探针”表示固定在固相支持物上的物质，而其他人把它理解为标记物。*Nature Genetics*（1999）的编辑鼓励《芯片预测》这篇综述的作者用“探针”表示被固定的核苷酸，而用“靶”代表游离的核苷酸（Phimister 1999）。尽管老式的 Northern 印迹法命名似乎要舒服一些，但我们还是采纳编辑的建议。本指南中，探针始终是固定的、未标记的和已知的，待测的靶永远是指标记的、在溶液中且未知的。

5 # 方案 1

cDNA 克隆集的复制和储存

Andrew Holloway，Maria Murphy（Peter MacCallum 癌症研究所，澳大利亚东墨尔本）；
Settara Chandrasekharappa（国立人类基因组研究所，国立卫生院，马里兰州贝塞斯达，20892）

本方案介绍了细菌宿主中克隆集的繁殖。克隆的重新纯化和序列核对大大减少了早期 UniGene 集中的误差（Halgren et al. 2001）。培养物的交叉污染可能会使这一努力白费，并导致克隆集基本上毫无用处。克隆集的污染通常是由于以下三种原因：不正确的复制技术、孔间培养物溢出、凝液从板盖滴入孔中。另一类型的污染由噬菌体 T1 感染造成。具有强攻击性的噬菌体能强烈抑制克隆生长，也大大限制了克隆集的价值。T1 对克隆集的污染就如同支原体对哺乳动物细胞系一样。检测和控制菌种中 T1 污染的方法将在关于**基因组和 cDNA 文库中噬菌体 T1 的防止和检测**信息栏中进行讨论。

第一次对克隆集进行操作时，最好核查培养物的无菌处理，如同本方案最初的步骤中所描述的那样。菌种的所有处理在层流的通风橱中进行，要特别小心以避免交叉污染。实验室有新的克隆集到达时要考虑的其他重要问题有：

- 需要保存多少拷贝？最少要制备一个工作拷贝和一个备份集。原始克隆集作为档案单独保存。
- 是否为原始的和拷贝的克隆集准备了充足的－80℃保存设施？一个 40 000 个克隆的集合，以 96 孔板形式运送，大约需 400 块板，约占一个标准－80℃冰箱的 1/3。
- 复制保藏物需要什么设备？为了保证克隆集保持有序和避免交叉污染有哪些注意事项？
- 克隆集的保存和使用由谁负责？许多人不受管理地接触克隆集将造成灾难。
- 对克隆集分发给其他研究者有无限制？大学和商业用户适用于不同的条件。

关于**克隆集**的信息栏中提供了一些常用克隆集及其衍生物的详细情况。

6 ## 材料

缓冲液和溶液

储存液、缓冲液和试剂的成分见附录 2。

将储存液稀释至合适的浓度。

羧苄青霉素（GIBCO）

羧苄青霉素储存液（1 g 羧苄青霉素溶于 10 ml 无菌水）

溶液用 0.22 μm 过滤器过滤，－20℃分装保存。

乙醇（95％和 70％）

据报道，某些批次的无水乙醇（100％或 200 标准强度）含有荧光污染物，在微阵列实验中产

生高的荧光背景，而95%乙醇不存在这样的问题。为避免这个难题，选用95%乙醇或确定一个品牌的无水乙醇不会产生高的荧光背景。购自Warner-Graham公司（Cockeysville，马里兰州）的无水乙醇（USP级乙醇，64-17-5）使用一直不错。

丙三醇水溶液（灭菌，45%和80%，*m/V*）

制备450 g/L酶级（enzyme-grade）丙三醇，高压蒸汽灭菌20 min，液体循环的压力15psi（1.05 kg/cm^2）（注：1 psi=6.894 76×10^3 Pa）。

次氯酸钠（3%）（漂白）

培养基

LB（Biofluids，Rockville，马里兰州）

含100 μg/ml羧苄青霉素的LB

将1 ml羧苄青霉素储存液加入1 L LB培养基中，未使用的含抗生素培养基保存不要超过1周。

含100 μg/ml羧苄青霉素/8%丙三醇的LB

将100 ml高压灭菌后降温至40℃以下的80%丙三醇加入到900 ml内含110 μg/ml羧苄青霉素的灭菌LB中。

载体和菌株

测试无菌技术的菌株

任何生长旺盛的大肠杆菌菌株均可。

待复制的克隆集

这些克隆集通常以菌种主平板（master plate）形式提供。

专用设备

微量滴定板黏性密封带（如Sealing Tape，7742000，ICN）

微量滴定板的冰箱架（PGC Scientific）

咨询－80℃冰箱供应商以选择合适的支架存放微量滴定板。

微量滴定板的热封膜(可去除)(如Easy Peel Film，AB-0559，Abgene，Surrey UK)

接种工具

可以得到一次性或可重复使用的不锈钢96针或384针工具（V&P Scientific，Genetix）。如果用自动设备，需使用带配套干燥装置的复制工具，如一种为Biomeck Robot（Beckman）设计的HDRT（高密度复制工具）。

层流的通风橱

液体操作工具

多通道移液器适合于克隆集的手工复制。许多种多通道移液器可供使用，从简单装置（可调式单次移液设备）到复杂的可编程机器，用于重复移取固定容量（供应商有Matrix，Oxford，Brinkmann）。

微孔板加样器，如Q-Fill微孔板加样器（Q-FILL2，GENETIX）一般是必需的。全自动的液体操作平台可从许多公司得到（见表1-2）。

微量滴定板（96孔，U型孔底；3799，Corning）

微量滴定板密封膜（Mylar透气膜；T640205，ICN）

微量滴定板密封膜（灭菌处理，96孔，如SEAL-THN-STR；Elkay产品，

Shrewsbury, Massachusetts)

培养液（4870，Costar）储存槽（50 ml，灭菌处理，一次性使用）

热封机（Combisealer，AB-0384/240，ABgene，Surrey UK）

方法

检验无菌技术和有效转移

在有价值的克隆集复制之前，非常重要的一点是确保子板的接种是可重复的，并保证复制工具在母板（source plate）间移动时进行了有效灭菌。在完全复制克隆集之前，如步骤1～5所述，要建立一个对照接种矩阵来仔细检查这两个参数。本方案针对96孔的形式设计，也可根据需要按比例放大到384孔的形式。

1. 在LB中制备10 ml标准的过夜细菌培养物。
2. 小心地移取150 μl过夜培养物至96孔板上轮流交替的孔中，余下的孔中加入不含抗生素的灭菌LB（图1-1，微孔板A）。在第二块多孔板上重复该过程，但是要互换含有培养物和未接种培养基孔的位置（图1-1，微孔板B）。

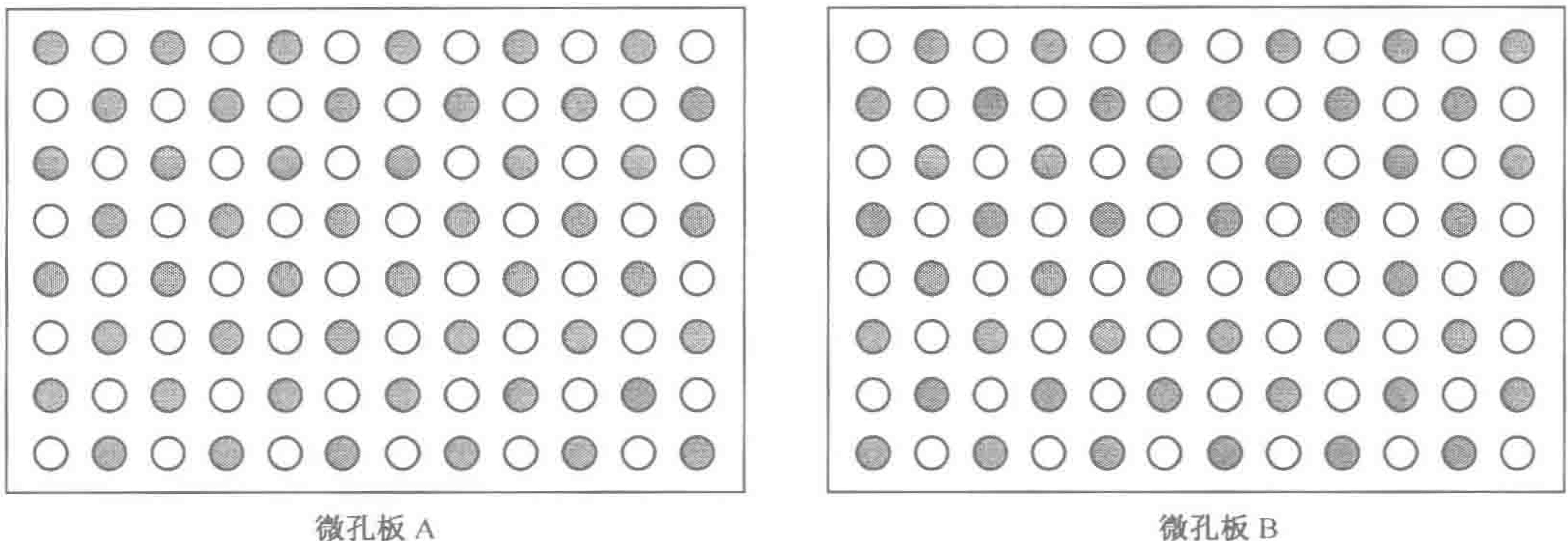


图1-1　96孔板的对照接种矩阵

微孔板A和B的接种样式相反。(实心环）接种有细菌培养物的孔；（空心环）带有灭菌接种培养基的孔

3. 向96孔板上的每个孔中加入150 μl不含抗生素的LB。
4. 用所选择的复制方法，将微孔板A复制成一个子板（板#1）。将复制工具灭菌或更换移液管吸头后，将微孔板B复制成第二个子板（板#2）。继续复制成对子板直到所有十块均被接种。

8

自动复制： 设定机器程序使得工具依次浸入3%漂白液、灭菌水和70%乙醇中各30 s。彻底干燥复制工具（通常需要鼓风干燥装置）。

手工复制： 准备一个内装80%乙醇的托盘，用于对接种针进行灭菌。将复制工具浸入乙醇，从乙醇中取出工具后对针头进行火焰消毒，使用前让复制工具稍微冷却（约10 s）。乙醇托盘需带有防火盖子。使用80%乙醇而不是无水乙醇，因为它可以减小发生火灾的可能性。

重要事项： 火焰消毒后勿将复制工具放回乙醇托盘。

5. 微孔板在37℃孵育过夜，不要振荡。

6. 将微孔板放在观片灯上进行检查。核对所有复制物的细菌生长样式与微孔板 A 或 B 一样。

   如果在应该无菌的孔内有细菌生长，要检查灭菌技术。大部分交叉污染是由于在移液操作中产生的气溶胶或者复制工具灭菌不充分引起的。为了避免这些问题，可采用防气溶胶的带滤芯吸头，并严格按上述实验方案对复制工具进行灭菌。

   一旦无菌技术和可靠的转移确立以后，就可以开始复制克隆集。

复制克隆集

7. 将密封的克隆集主板在 30℃下孵育过夜。

   供应商提供的细菌培养物有时密度较低。直接用这些稀释的菌种复制可能导致子培养物无法生长。

   培养物在 30℃而非 37℃生长避免了耗尽抗生素，这会在带有多拷贝质粒的细菌生长达到饱和时发生。保持充足的抗生素浓度对避免无质粒细菌的生长很重要。

8. 标记一组圆 U 型底部的标准 96 孔板，并向每个孔中加入 100 μl 含 100 μg/ml 羧苄青霉素/8%丙三醇的 LB（384 孔板加入 50 μl/孔）。微孔板底部和顶部均作标记。

   这些微孔板将用作工作拷贝。将培养基倒入灭菌的一次性储存槽，并用 8 或 12 通道移液器和自动设备，或者用配备 8 通道适合于 96 孔板，或者配备 16 道适合于 384 孔板的 Q-Fill 装置，将培养基转移到微孔板上的孔中。

   检查所提供的克隆集信息，确保克隆的 cDNA 被携带于带有氨苄青霉素抗性的载体中，这种情况通常存在于人 IMAGE 克隆中。羧苄青霉素是一种半合成碳氧青霉素，比氨苄青霉素更稳定，因此它是带有质粒的克隆推荐的选择试剂。

   丙三醇（最终浓度 8%）在培养基的冷冻过程中用作冷冻保护剂，在培养物生长前后均可加入。但还是建议在生长前加入，因为这样可以避免进一步的操作步骤，从而减少污染的机会。

9. 将主板放在水平的微量滴定板转头上，室温下 1300 *g* 离心 2 min，以便在开启前将密封膜上的凝液和水滴除掉。

   在密封膜内表面的细菌培养液很容易在孔间转移，造成菌种交叉污染。

   如果从冷冻的微孔板开始操作，则无需离心微孔板；当孔内物质仍处在冷冻状态时，除去密封膜，然后让培养液在室温下解冻（请见本方案结束部分关于**附加方案：使用冷冻菌种**的信息栏）。

10. 将每块子板（待接种的）置于相应的主板旁，使得两者的行和列都按同样的顺序排列。将复制工具浸入主板，然后再浸入子板。

    如果主板和子板不按照同样的方式定位，可能会引入误差。当主板和子板叠放在一起时，在微孔板角落接近 A-1 孔处用黑色毡尖笔从上向下画线标记。当微孔板分开时，该记号可使它们易于定位。给每块微孔板和盖子编上号码。

    选用平头而非尖头的 96 针工具移取足量的细胞。另外，将 96 针工具浸入直至与孔底接触，然后用针轻轻搅动培养物确保移取足量的接种物。这一步使用自动装置能大大提高接种效率。转移克隆不但费力费时，而且容易受到操作者误差影响。使用自动装置可以减少错误转移的风险。

    如果要复制子板的多个拷贝，每轮复制要给针重新接种。

11. 最好用透气的 Mylar 微孔板密封膜密封接种好的子板，或者更换微孔板盖。确

保每次接种间隔对接种工具进行灭菌。作为对照，每次实验中包含一块用灭过菌的工具接种的空白微孔板和一块没有接种的空白微孔板。

作为首选，微孔板要用透气的密封膜进行密封，这使得在减少污染机会的情况下有足够的气体交换来维持生长。简单地更换微孔板盖子将增加交叉污染的机会。透气密封膜不适合于在－80℃下保存，在生长结束后和保存前需换成合适的耐低温密封膜。

12. 将密封或者加盖后的接过种的 LB 平板置于孵育箱中，30℃生长过夜，不振荡。

培养基的浑浊（和细胞在底部沉积）表明细胞在生长。为了达到合适的密度，微孔板可能需要孵育 18～20 h。

不需要振荡培养物。为避免培养液失水干燥，将微孔板置于潮湿袋（humidified bag）中，如 1 L“拉链”袋，其中含有潮湿的纸巾。复制板生长以后，要检查空白微孔板是否保持无菌。

13. 如果生长培养基中不含丙三醇，向每个孔中加入 50 μl 45％（m/V）的灭菌丙三醇。用气密的密封膜密封微孔板，更换板盖，堆放在保存架上。将微孔板在－80℃保存。

10

克隆注释工作表

	母板			目的板		
克隆 ID	板	行	列	板	行	列
AA324398	1	H	1	1	A	1
AA755111	2	E	1	1	A	2
AA947096	3	D	3	1	A	3
AA430733	6	G	12	1	A	4
AA229529	7	D	5	1	A	5
AA203631	9	E	6	1	A	6
AA562669	15	A	8	1	A	7
AA972704	22	E	3	1	A	8
AA925652	22	D	11	1	A	9
AA828055	26	B	10	1	A	10
AA937131	40	G	7	1	A	11
AA635869	40	H	8	1	A	12
AA870438	43	C	2	1	B	1
	⋮	⋮	⋮	⋮	⋮	⋮

11

附加方案：使用冷冻菌种

一个成功的微阵列实验要求原始的冷冻菌种平板能用于恢复感兴趣的 cDNA 插入片段以进行下一步分析。来自菌种母板的克隆可扩展为在新鲜微量滴定板孔内或 15 ml 试管中生长的细菌培养物。无论用哪种方式，在任何情况下重新分离单个细菌克隆进行下一步分析和测序都必须谨慎。下面的实验方案介绍了手工选取克隆的操作。有几种自动克隆选取设备可供使用，使用方法参照生产商提供的说明书。

1. 准备一个克隆注释工作表（见前页），列出微孔板的编号和冷冻克隆集中的克隆的位置。准备一块接种板，每个目标孔中含 100 μl 合适的培养基。

另外，也可以准备一组含1～2 ml培养基的15 ml试管。如下面例子中所介绍的，用克隆的唯一标识符及其位置（即板和孔的编号）清楚地标记试管。

重要事项：在克隆选取过程中保持无菌技术。实验过程中要包括对照以检查生长培养基和选取设备的无菌情况。

2. 用复制的菌种平板而不是主板进行实验。最好将平板放在一个有光照的盒子中，下方放有模板，标记行和列的号码。
3. 培养液仍处在冷冻状态时，依次去除每块平板的密封膜。用一根灭菌的牙签或移液器吸头轻刮合适克隆的冷冻表面，并接种目标培养基。取样后在每个孔盖上标记一个"十"字，用来保持已取过样孔的记录。

   重要事项：不能让平板完全解冻。反复冻融会影响活力。使平板升温到菌种稍微融解可使克隆选取任务易于进行，能方便地刮取其表面。

4. 用防冻密封膜或黏性密封膜重新密封好提供菌种的平板（donor plate），放回冰箱。

   这步要迅速操作，以免热封机使培养基升温过高。

5. 将复制培养物孵育过夜。对于微量滴定板培养物，生长条件与复制原始克隆集一致（主要方案的步骤12）。15 ml试管中的培养物要在37℃孵育，并充分振荡。一旦过夜培养物生长后，在含100 μg/ml羧苄青霉素的LB琼脂平板上划线培养得到单个克隆。
6. 选取4～6个分离较好的克隆并使它们过夜生长。从它们和父代培养物中制备质粒DNA（步骤5）（方案2、可选择方案或Sambrook and Russell 2001）

   这些克隆中的cDNA插入片段要通过限制性消化切除并用琼脂糖凝胶电泳检测，以确保在父代培养物中存在单一条带，并且相同尺寸的条带在次培养菌（subculture）中存在。

7. 对从次培养菌质粒分离的cDNA插入片段进行测序，并核对DNA序列与跟随克隆集一起提供的工作表中序列一致。

   在使用一个克隆集之前，一些研究者随机选择克隆重新测序以测量克隆集中错误注释的原始比率。

   保留克隆集使用者的记录并汇总他们的经验。这样可以积累关于再次污染（de novo contamination）水平和（或）复制的克隆集中错误注释的信息。

方案2

来自细菌克隆集探针的PCR扩增

Andrew Holloway, Maria Murphy（Peter MacCallum癌症研究所，澳大利亚东墨尔本）；
Settara Chandrasekharappa（国立人类基因组研究所，国立卫生院，马里兰州贝塞斯达20892）

制备cDNA点样微阵列中最费力的是通过高通量PCR制备用于阵列的cDNA片段的浓缩液。其目的是从大部分克隆中扩增插入片段，避免交叉污染，并从酶、核苷酸、引物和其他PCR组分中纯化片段。尽管这一过程很费时间，但如果操作正确的话，每

100 μl PCR 中就可以得到足够的材料，点样数千片微阵列玻片。

因为需要大量的扩增片段，所以希望操作过程能结合一定程度的自动化，如 Multimek 或 Biomek FX（Beckman）多通道自动移液设备。自动化仪器不仅可以增加产量，而且能够减少由于人工操作带来误差的几率。在没有自动设备的情况下，对包含几千个克隆的克隆集进行扩增是富有挑战性的。液体操作设备的清单请见本方案结尾部分的表 1-2。

在考虑克隆集和载体/引物对时，参考本章导言部分的指南和关于**克隆集**的信息栏。大部分克隆集提供了扩增的说明。但在使用供应商的方案进行大规模扩增前，建议使用一块克隆平板进行预备实验。

下面方案是为 96 孔板形式设计的，按相应比例放大后可适用于 384 孔板形式。

材料

缓冲液和溶液

储存液、缓冲液和试剂的成分见附录 2。

将储存液稀释至合适浓度。

10×扩增缓冲液（由 Applied Biosystems 提供，N808-0190）。

或者按以下配方制备 10×扩增缓冲液效果良好：

50 mmol/L KCl

100 mmol/L Tris-HCl（pH8.3）

将缓冲液分装，在－20℃保存

H_2O（灭菌的）

Milli-Q 水有时 pH 很低。也可选择使用商品注射用水（Baxter Healthcare）。

$MgCl_2$（25 mmol/L）

酶和缓冲液

Taq DNA 聚合酶（5 单位/μl）（如由 Applied Biosystems 提供的 Amplitaq Gold）

许多实验室中使用由表达该蛋白质的重组体大肠杆菌中提纯的 *Taq* 聚合酶。Pluthero（1993）给出了简单的纯化策略。

核酸和寡核苷酸

扩增引物

常用克隆集插入片段的扩增引物见表 1-1。

表 1-1 扩增用引物

载 体	使用的文库	正向和反向引物
pT3-T7Pac	Soares	正向：ctgcaaggcgattaagttgggtaac
		反向：gtgagcggataacaatttcacacaggaaaacagc
pCMV-Sport 6	MCI-GAP	正向：ctgcaaggcgattaagttgggtaac
		反向：gtgagcggataacaatttcacacaggaaaacagc

续表

载　体	使用的文库	正向和反向引物
pBluescriptSK-	NCI-GAP；Stratagene	正向：ctgcaaggcgattaagttgggtaac
		反向：gtgagcggataacaatttcacacaggaaacagc
LafmBA	Soares 1 NIB	正向：ctgcaaggcgattaagttgggtaac
		反向：gtgagcggataacaatttcacacaggaaacagc
pCMV-Sport	Beddington embryonic region	正向：ggagagagctatgacgtcgc
		反向：gacactatagaaggtacgcctgc
pCMV-Sport 1	NIA	正向：ccagtcacgacgttgtaaaacgac
		反向：gtgtggaattgtgagcggataacaa
pCMV-Sport 2	Life Tech embryo	正向：ggagagagctatgacgtcgc
		反向：gacactatagaaggtacgcctgc
pOT2A	BDGP *Drosophila* library	正向：aatgcaggttaacctggcttatcg
		反向：aacgcggctacaattaatacataacc
带有 M13 引物位点的载体[a]	广泛应用多种文库	正向：gttttcccagtcacgacgttg
		反向：tgagcggataacaatttcacacag

a 推荐 Hegde 等（2000）描述的引物以避免某些载体通用引物位点中存在的点突变。

反义引物（100 μmol/L）水溶液

将储存液分装，在－20℃保存。

dNTP 溶液，含全部四种 dNTP，浓度均为 10 mmol/L。

这些溶液可从，如 Amersham Pharmacia 购买，或按附录 2 所述制备。将储存液分装，在－20℃保存。

正义引物（100 μmol/L）水溶液

将储存液分装，在－20℃保存。

专用设备

带滤芯吸头

带适配器的台式/医用离心机

这种离心机能离心深的和标准尺寸的 96 孔微量滴定板（如 Eppendorf 5810R、来自 Beckman Coulter 的 Allegra 6R 或是其他相同的）

Microseal "A" film 密封膜（MJ Research）

微量滴定板（薄壁的，100 μl）

多通道移液器或自动液体操作设备

带加热盖的热循环仪器［如可容纳 4×96 孔板的 Tetrad（MJ Research）或 Gene-Amp 9700（Applied Biosystems)］

方法



模板的制备

1. 根据方案 1 复制平板，制备克隆的新鲜过夜培养物以备扩增。

本方案中，用细菌细胞的新鲜过夜培养物作为模板，从热灭活的细菌直接扩增。用新鲜培

养物作为模板 DNA 来源可大大提高 PCR 产物的产量。定好计划，可将克隆集的复制和 PCR 扩增步骤连接起来。只需复制与 PCR 中所使用的量一样多的平板。使用每种培养物的一部分生成 PCR 模板（见步骤 2）并将剩余部分冷冻。培养物中的丙三醇不会影响后续的 PCR 扩增。

2. 向薄壁 PCR 微量滴定板的每个孔中加入 90 μl 灭菌水。
3. 用多道移液器或自动液体操作设备混合过夜培养物（方案 1 制备）

   必须将细菌培养物重新悬浮以获得足量的 PCR 模板。
4. 用多道移液器或自动液体操作设备移取每种新鲜过夜培养物各 10μl 到加水的微量滴定板孔中（方案 2 中准备）。

   使用带滤芯的吸头以减少交叉污染的危险。
5. 用黏性密封带封好微量滴定板后放入 100℃热循环仪中 10 min。
6. 1300 *g* 下将微孔板离心 5 min，可见到细胞碎屑的小团块。
7. 每次 PCR 扩增以 1 μl 上清液为模板，平板在－80℃下保存。储存液可反复冻融数次，而不会影响 PCR 产物量。

   在 PCR 中可以直接用 1 μl 新鲜过夜培养物作为模板，并在第一次 PCR 循环中将变性步骤的时间延长至 5 min，在原位溶解细菌（Yue 2001）。这种方法要迅速一些，但产量较低，且失败率增加。

   经过纯化的 PCR 产物可用作下一步 PCR 扩增的模板。反应产物用 TE（pH7.9）或水重悬，浓度为 100 ng/ μl。用 1 μl 稀释后的 DNA 作为模板。

   如果失败率过高（＞5%）或产率不稳定，考虑使用纯化后的质粒作模板，取代离心的裂解物（见本方案结尾部分关于**可选方案：作为扩增模板的质粒 DNA 的制备**信息栏）

cDNA 克隆的扩增

表 1-1 介绍了扩增的常用引物。下面一系列实验步骤提供了通用的指导，适合使用 M13 正向和反向引物的细菌或质粒模板。

15

8. 对每块 96 孔板，将下列试剂在冰上混合配制扩增反应混合液（PCR 混合液）：

试剂	用量
10×扩增缓冲液	1.1 ml
25 mmol/L $MgCl_2$	1.1 ml
4 种 dNTP 的 10 mmol/L 溶液	220 μl
100 μmol/L 正义引物	55 μl
100 μmol/L 反义引物	55 μl
5U/μl *Taq* 聚合酶	55 μl
H_2O	8.3 ml

   这里的制备物足够进行约 105 次 PCR，考虑到细微的移液误差和其他损耗，提供过量的混合物。
9. 给适当数量的灭过菌的薄壁 PCR 微量滴定板做好标记，写上日期。
10. 向每个孔中加入 99 μl PCR 混合液。

    确定供给板（含细菌模板）和接受板（recipient plate）（含 PCR 混合液）之间定位是否正确。
11. 用多道移液器或自动液体操作设备移取 1 μl 细菌裂解物模板至每个孔中。

每行加入模板后更换吸头，再次检查供给板和接受板定位是否正确。

12. 根据制造商的说明用硅密封膜封好平板。

13. 将微量滴定板放入热循环仪中，按照下列变性、复性、聚合时间和温度扩增DNA模板。

循环次数	变性	复性	聚合
1	95℃ 60s		
30	94℃ 30s	65℃ 45s	3 min，72℃ 30s
最后一次循环		72℃ 5min	

这些时间条件针对本方案介绍的专用设备和反应是最佳的，但对其他类型的设备和反应，还需要调整。例如，当直接从（未裂解）细菌扩增时，最初的变性实验步骤需增加至5min。一般推荐30次循环，一些克隆集可能需要更长的循环次数，如国家衰老研究所（NIA）的小鼠15K集建议使用38次循环。

14. 循环完成后，将微量滴定板从热循环仪中取出，为点样微阵列而继续纯化和分析（见方案3和方案4）。

作为替代，平板可在4℃保存24～28 h，或在－20℃保存更长时间。尼龙滤膜微阵列不需要对PCR产物进一步纯化（请见第1章，方案1）。

表 1-2　液体操作

公　司	平　台	评　价
Beckman (http://www.beckman.com)	Biomek 2000	单道和八道液体操作仪(单头并带有八种可更换的器具,用于移液及稀释);可使用不带滤芯或带滤芯的吸头 移液体积从1 μl～1 ml;工作台上可同时放置4～8块微孔板。该系统可升级成为高通量,要装配1～2件Stacker Carousel装板器，每件Stacker Carousel容纳120块微孔板
	Multimek-96	96孔自动液体操作仪,具备高通量试剂加入和转移 移液体积从1 μl～1 ml。系统可在同一工作台上应用96-96或96-384模式
	Biomek FX	配有移液头的单或双桥系统,包括Span8、96或384规格 移液体积从0.5～200 μl，使用Span8系统可实现大体积液体分配 该系统可配备1～2个Stacker Carousel而升级到高通量,每个Stacker Carousel容纳120块微孔板
Robbins Scientific (http://www.robsci.com)	Hydra 96	系统为96孔板移液而设计,配备有100 μl、290 μl、580 μl和1 ml注射器。使用100 μl注射器时,移液体积最低为100 nl。升级可选择自动洗板系统、微孔板定位平台和装板器
	Hydra-384	系统为384孔板和其他高密度多孔板设计,配备有100 μl、290 μl、580 μl和1 ml注射器。使用100 μl注射器时,移液体积最低可至100 nl。升级可选择自动洗板系统,微孔板定位平台和装板器
	Tango Liquid Handling System	可配置1、8、12、96和384道的可拆卸式分液头,配有100 μl、290 μl、580 μl和1 ml注射器。使用100 μl注射器时,移液体积最低可至100nl。系统配有12-工位XY平台以及平台可选的能容纳240块微孔板的装板器

续表

公　司	平　台	评　价
Genetix (http://www.genetix.com)	QBot	QBot 可选配置 96 道液体操作头，使用一次性塑料吸头。系统自动装卸吸头 系统吸取液体和分液体积范围 1～50 μl 与 96 或 384 孔板兼容 载板量：母板 72 块，目标板 24 块
	QPet	独立的 96 或 384 道吸取液体和分液工作台系统 该系统与 96 和 384 孔板兼容，能够装配一次性塑料吸头或不锈钢针。有两种规格吸头，容积为 1～30 μl 或 10～300 μl。不锈钢针的容积为 0.5～100 μl 载板量：母板或目标板——100 块标准微孔板或 30 块深孔板
BioRobotics (http://www.biorobotics.com)	MicroGrid Ⅱ	装配多根针的工具，每次可在母板和子板之间最多转移 384 个样品。扩展和(或)压缩模式下，可以在相同密度或不同密度孔板之间进行转移 移液体积在 nl 范围 载板量：最多可载 48 块微量滴定板 该系统可以处理 1536 孔板
Qiagen (http://www.qiagen.com)	BioRobot 3000	系统有 96、384、96—384 或 384—96 规格 移液体积从 1 μl～2.5 ml，使用 4 根针(不锈钢针或一次性吸头，或者它们的组合) 使用装板器可升级为高通量
	BioRobot 8000	系统有 96、384、96—384 或 384—96 规格 移液体积从 1 μl～2.5 ml，使用 8 根针(不锈钢针或一次性吸头，或者它们的组合) 使用装板器可升级为高通量
	BioRobot Rapid Plate	系统有 96、384、96—384 或 384—96 规格 移液体积从 1 μl～2.5 ml，使用 96 根针(一次性吸头) 使用装板器可升级为高通量
Tecan (http://www.tecan-us.com)	Genesis Robotic Sample Processor	装配 4～8 支吸头的操作臂实现高通量。同支操作臂上的吸头可以是标准的和一次性吸头的任意组合 使用 200 μl 或 1000 μl 一次性吸头，带有或不含滤芯，移液体积从 1 μl～1 ml 与 Genesis 工作站集成，实现中等通量操作(200 块微孔板)
Packard BioScience (http://www.packardbio science.com)	MiniTrak MPD System (Ⅳ, Ⅴ, Ⅵ, Ⅶ)	小型液体操作模块，可配置成包含 4～6 个模块的系统，用于为 96/384/1536 孔板移液 包含 96 个一次性吸头的 MPD(多位置分配模块)移液头，移液体积为 5～235 μl；分配器可以定位，对 96、384 和(或)1536 孔板吸取和分配液体 模块可升级至 96 吸头的 P50 型分配头(容积范围 0.5～50 μl)或 384 吸头 P30 型分配头(体积范围 0.5～30 μl) 使用 Universal Stacker Modules 可以将系统升级为更高通量；使用装板器，每个模块能容纳 50 块 SBS 标准微孔板，15 块深孔板或 PCR 板

续表

公司	平台	评价
	MultiProbe Ⅱ	MultiProbe Ⅱ系列产品提供多种尺寸的平台 标准系统提供的平台可容纳 16 个微孔板架(最多可放置 16 块 96 孔板) 扩展系统设计成中央部分能容纳 16 个微孔板架,但还可以增加两个扩展模块,每个模块能容纳 8 个微孔板架。(最大容量为 32 块 96 孔板) 4 个独立的吸头,可以是一次性的或者固定吸头(固定吸头是可清洗的,覆盖了聚四氟乙烯的不锈钢取样针)。可选的超低容积吸头移液从 200 nl 到数毫升。使用 20 μl、200 μl 或 1 ml 的吸头,标准吸头移液体积从 1 μl～1 ml
	MultiProbe Ⅱ HT	8 个独立的吸头,包括一次性的或者固定吸头(固定吸头是可清洗的,覆盖了聚四氟乙烯的不锈钢取样针)。超低容积吸头的移液体积从 500 nl 到数微升
Hamilton (http://www.hamilton comp.com)	MicroLab 4000 Series	MicroLab 4000 提供了可容纳 18 块微孔板的平台和一个主动式清洗工作站 平台容量 MPH-96 型:纵向排列时 34 块微孔板,横向排列时 38 块微孔板 MPH-48 型:纵向排列时 16 块微孔板,横向排列时 18 块微孔板 MicroLab 4200 的扩展平台提供了可容纳 38 块微孔板的平台和一个主动式清洗工作站 MPH-96 的平台容量更大,配备了 96 吸头的装置,而 MPH-48 的平台要小一些,配备 48 吸头的装置。两者具有相同的性能和指标 吸头装置能够在横向或纵向上到达平台的所有位置,包括自定义的微孔板位置;能够同时进行 96 或 48 道移液,或者对 8 个注射器的分组进行独立控制,实现在平台的纵向上对单行和横向上对单列的控制
MWG (http://mwgbiotech.com)	RoboSeq 4200	单道和 4 道液体操作仪 可清洗的和(或)一次性吸头。移液体积最低 1 μl。能够工作在 96/384 孔方式和多种试管方式 特点有 Peltier 制冷的试剂架,移液工作站及装板器 可选配置:集成的 MWG 热循环器,微孔板读板器,真空装置;DNA 测序梳状加样器
	RoboSeq 2500	8 道和 16 道液体操作仪 可清洗的和(或)一次性吸头。移液体积最低 1 μl。能够工作在 96/384 孔模式 特点有 Peltier 制冷的试剂架,移液工作站及装板器 可选配置:集成 MWG 热循环器,真空装置
	RoboSmart	为匹配毛细管 DNA 测序仪设计的 8 通道系统。提供各种部件的简易式组合:质粒制备,序列测定,集成循环,染料标记终止物去除
Lawrence Berkeley National Lab Human Genome Center (http://www.jgi.doe.gov/programs/instrumentation/preptrack.htm	Prep Track Robot	系统可配置多种可能的模块 高容积分液工作站有 4 个 96 通道移液模块(每个模块配备 96 支带有固定吸头的注射器) 移液体积从 5 μl 到数毫升 系统传入和传出处的微孔板运输模块能够通过传送带将微量板送入或运出,实现连续操作

可选方案：作为扩增模板的质粒DNA的制备

有些实验室发现，使用纯化的质粒代替离心的裂解物进行扩增反应，可以提高扩增产率。质粒DNA可以用1～1.3 ml的细菌培养液制备，使用市售的小量制备试剂盒，根据制造商的使用指南进行操作即可。

附加材料

注意事项：操作带〈!〉标记的材料时，请参考附录3。

琼脂糖凝胶（1%），放入1×TAE中，含有0.5 μg/ml溴化乙啶〈!〉

培养皿（如Qiagen 19579平底培养皿，或者Millipore有盖细胞培养皿LSKC CB500）

质粒小量制备试剂盒［如Millipore Montage Plasmid Miniprep 96试剂盒（LSKP09601）或Qiagen Qiaprep 96 Turbo Miniprep试剂盒（27191）］

复制器（96针）或带有96针复制工具的自动液体操作仪

TE（pH 8.0）

10 mmol/L Tris-HCl（pH 8.0）

0.1 mmol/L EDTA

含适当浓度抗生素的Terrific Broth培养基

方法

1. 用多道移液器或自动设备向96深孔培养皿的每个孔中加入1～1.3 ml含适当抗生素的Terrific Broth培养基。
2. 用96针复制器或带有96针复制工具的自动液体操作仪为培养皿的每个孔接种细菌。

   注意使供样（donor）板、接受（host）板排列方向相同。另外，如果使用自动设备，要保证复制工具与孔中的生长培养基接触。为了避免交叉污染，在不同板间进行移动时，要对复制器进行彻底的清洁。
3. 用透气的密封带密封培养皿，置于速度为200～300 r/min的摇床上，37℃下培养16～22 h。
4. 使用市售的小量制备试剂盒，根据制造商的使用指南进行操作，从过夜的培养物中分离得到质粒DNA。

   该步骤通常包括，用碱裂解细菌，再通过基于树脂的亲和净化系统洗出纯化的质粒DNA。
5. 通过琼脂糖凝胶电脉、溴化乙啶染色来估测有代表性的纯化质粒DNA样品中DNA的浓度。为了进行定量，把1 μl样品与已知浓度（50 ng、100 ng、150 ng）的超螺旋质粒DNA标样作比较。
6. 把质粒DNA重新悬浮到TE（pH 8.0）中，最终浓度为100 ng/μl。

   在这一阶段，DNA制备物放在小量制备试剂盒内提供的96孔板的孔中。

重悬溶液的体积随质粒 DNA 的产率而变化，而产率又依赖于载体和宿主株。对于绝大多数载体/宿主株组合，标准的产率为 20 μg/孔。

7. 用 1 μl 的质粒 DNA 制备物作为 PCR 扩增的模板（主要实验方案中的步骤 12）。剩余的质粒 DNA 在－20℃下储存。

方案 3

PCR 产物的纯化

Andrew Holloway，Maria Murphy（Peter MacCallum 癌症研究所，东墨尔本，澳大利亚）；
Settara Chandrasekharappa（国立人类基因组研究所，国立卫生院，贝塞斯达，马里兰州 20892）

微阵列点样前对 PCR 产物进行纯化，可以除去导致打印头堵塞的酶或者其他蛋白质，以及与产物竞争结合底物的寡核苷酸引物，并且可以除去影响与底物结合的盐类。这里介绍了两种纯化方法：尺寸排阻（size exclusion）法和乙醇沉淀法。基于尺寸排阻的方法较好，但是成本较高。

这里介绍的尺寸排阻法使用 Millipore Multiscreen 板，还有其他的类似产品（如 CentriSep-96，Princeton sepetration）。乙醇沉淀法将在本方案结尾部分关于**可选方案：用乙醇沉淀法纯化扩增产物**的信息栏中介绍。

材料

缓冲液和溶液

储存液、缓冲液和试剂的成分见附录 2。
将储存液稀释至合适浓度。

灭菌水

点样缓冲液（见方案 5）

核酸和寡核苷酸

扩增产物（见方案 2）

专用设备

台式/医用离心机，带有适配器。

这些离心机要能够离心深孔和标准的 96 孔微量滴定板（如 Eppendorf 5810R、来自 Beckman Coulter 的 Allegra 6R 或是其他相同的）。

微量滴定板摇床（MPS1，Ratek Platform Shaker）

多道移液器

Multiscreen PCR 板（MAN U03010，Millipore）

Multiscreen 真空集管装置（vacuum manifold）（MAVM096 OR，Millipore）

旋转真空浓缩器，带有微量滴定板适配器

21 方法

1. 用多道移液器将扩增反应产物（见方案 2）转移到 Multiscreen PCR 板上。将板置于真空集管装置上，在 67.73 kPa 真空下处理 5～10 min，直至板孔中没有残留液体。
2. 洗板两次，每个孔用 50 μl 水。按照步骤 1 方法，真空处理除去清洗用水。
3. 把样品重新悬浮于 100 μl 的灭菌水中。将板置于剧烈振荡的平台上 10 min，确保板孔中的液体成漩涡状。
4. 从摇床上取下板，使用多通道移液器或自动设备把样品转移至新的 96 孔或 384 孔微量滴定板中。
5. 每个样品中取出 0.5～2 μl，进行凝胶电脉分析（见方案 4）。
6. 在真空浓缩器中干燥剩余的样品，点样前将板置于－20℃下储存。关于重悬样品准备点样，见方案 5。

   使用旋转真空浓缩器和适当的真空压［15 mbar（1 bar＝10^5 Pa）］来干燥样品很重要，这样 DNA 既不会冻干，也不会完全脱水。真空处理时采用除去全部可见痕量水溶液所需要的最短时间。干燥之后，最好让所有的 DNA 停留在孔底的最小区域内。如果样品是冻干的，那么 DNA 往往沉积在板的侧壁，当样品重新水化时就有可能不会再次溶解。

22 可选方案：用乙醇沉淀法纯化扩增产物

扩增产物可以从反应混合物中回收，并在乙醇存在条件下通过沉淀法进行纯化。

附加材料

乙醇（95％和 70％）

据报道，某些批次的无水乙醇（100％或 200 标准强度）含有荧光污染物，在微阵列实验中产生高的荧光背景，而 95％乙醇不存在这样的问题。为避免这个难题，选用 95％乙醇或确定一个品牌的无水乙醇不会产生高的荧光背景。购自 Warner-Graham 公司（Cockeysville，马里兰州）的无水乙醇（USP 级乙醇，64-17-5）使用一直不错。

微量滴定板（V-型孔底，聚苯乙烯），可以在 1300 *g* 下离心（如 3894，Corning）

醋酸钠（3 mol/L，pH 5.2）

方法

1. 将扩增产物转移至 V 型底部的聚苯乙烯微量滴定板。使用多道移液器向每个板孔中加入 5 μl 的 3 mol/L 醋酸钠（pH 5.2）和 150 μl 的 95％乙醇。

   薄壁的 PCR 板在离心过程中可能会形变。不能使用聚丙烯板，因为 DNA 不能黏附到其表面上。
2. 在－80℃下沉淀 DNA 1 h，或在－20℃下过夜。
3. 4℃，1300 *g* 离心 1 h，回收 DNA。

4. 小心地倒出上层清液，把微孔板倒扣在吸水纸上沥干。用多道移液器沿着板上每个孔的侧壁各加入 150 μl 的 70%乙醇，注意不要移动 DNA。
5. 4℃，1300 *g* 离心 1 h，回收 DNA。
6. 按照步骤 4 所述除去上层清液。让微孔板在层流通风橱里干燥过夜。
   乙醇沉淀后过分干燥会导致 DNA 难以重新悬浮。
7. 给微量滴定板盖好盖子，用 Saran Wrap 密封，于－20℃下储存。也可以把 DNA 重新悬浮于 100 μl 水中，用于分析和点样（关于样品重新悬浮用于点样，参见方案 5）。

方案 4

PCR 产物的观察和评估

Andrew Holloway，Maria Murphy（Peter MacCallum 癌症研究所，东墨尔本，澳大利亚）；
Settara Chandrasekharappa（国立人类基因组研究所，国立卫生院，贝塞斯达，马里兰州，20892）；
Damares Monte（华盛顿卡内基研究院，斯坦福，加利福尼亚，94035）

DNA 产率、存在游离引物或引物二聚体，以及每个反应的电泳条带数目是衡量用于点样 PCR 扩增材料质量的重要指标。质量控制可以在 PCR 产物纯化前或纯化后进行。在纯化前进行检测分析可以避免由于处理失败的 PCR 产物而带来的人力和物力的浪费。另一方面，在纯化后进行检测分析则可以更好地估算最终产率。我们推荐在纯化前进行试验性的扩增反应，检测扩增的产率和一致性。一旦得到了可靠的实验结果，就最好是在样品纯化后检测其质量。

对 PCR 产物的定性及定量分析可采用琼脂糖凝胶电脉（在本实验方案中介绍）与 PicoGreen 荧光染料染色（在关于**附加方案：用 PicoGreen 进行定量**的信息栏中介绍）两者相结合的方法。按照如下所述的方法制备上样板，可以有效且连贯地进行这两种分析。

制备任何容量的琼脂糖凝胶来分析大量的 PCR 产物是很容易的。通过多道移液器进行凝胶上样时，必须使用具有适当孔间距的凝胶梳。可获得的一些商用系统（如现成的或快速分析凝胶电泳）能够提供快速有效的分析，但分辨率有限。首轮 PCR 循环的错误率（多条带或无产物）应该控制在 5%以下，这些最初的错误大约有一半可以在随后的一次 PCR 循环中得到挽回。

材料

注意：在操作〈!〉标记的材料时，请参考附录 3。

缓冲液及溶液

储存液、缓冲液和试剂的成分见附录 2。
将储存液稀释至合适浓度。

6×凝胶上样缓冲液［用溴酚蓝/二甲苯蓝（xylene cyanol）示踪染料］〈!〉
水（灭菌）
TE（pH 7.5）

24 **核酸和寡核苷酸**

DNA 分子质量标记序列梯（DNA ladder）（1kb）（15615-016，GIBCO）
待测样品［PCR 扩增的直接产物（方案 2）或纯化产品（方案 3）］

凝胶

预制商用凝胶体系（如即用型）（1.2%琼脂糖 96 孔凝胶，80-6461-33，Amersham）
电泳系统（80-6460-95，Amersham）

专用设备

铝箔带（3M Scotch）
微量滴定板（96 孔，V 型底部的聚丙烯制品，如 MLP-9601，MJ Research）
多道移液器

方法

凝胶板（gel plate）的准备

1. 新的 96 孔板中，每个孔带有 0.4 μl 6×凝胶上样缓冲液，再向每个孔中加入 2 μl

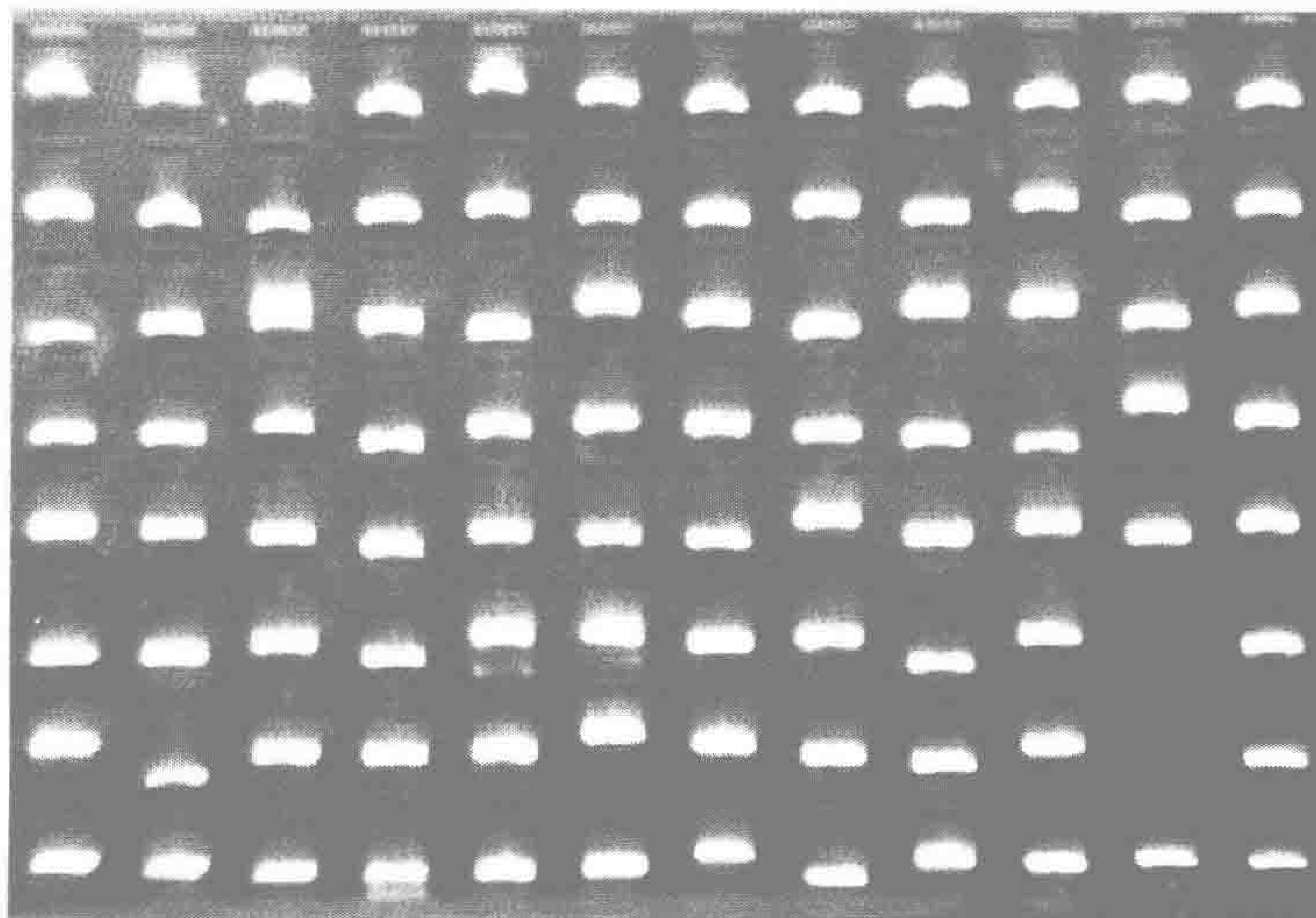


图 1-2　96 孔板中 PCR 产物的分析

溴化乙啶染色的 DNA 用 1.2%即用型凝胶电泳法分析，并在紫外光照射下观察。在本例中，扩增反应进行良好，因为大多数泳道中每种 DNA 均有一致的高产率，仅出现两个不成功的反应

PCR 产物或是经过纯化后的产物。通过吸液操作把样品混合均匀。用 3M Scotch 铝箔带覆盖微孔板，以防溶液蒸发。如果不在当日进行凝胶操作，则要把板保存在－20℃下。

2. 从凝胶板上移取每种 DNA 样品溶液各 2 μl，上样到即用型凝胶的每个孔中进行电泳，直到溴酚蓝迁移到大约为凝胶长度一半时为止。

   可用 250 ng 1 kb 序列梯作为分子质量标记物，以帮助进行定量。

3. 在紫外光照射下进行凝胶成像。关于经过染色凝胶的例子，见图 1-2。

   由于即用型凝胶中已含有溴化乙啶，所以电泳之后不再需要进行凝胶染色。

凝胶评分

4. 按照下述方法对 PCR 产物进行质量评分（关于每个分值代码的示例，见图 1-3）。

0	合格	样品只有一条具有高强度的条带。
1	空白	凝胶中观察不到 PCR 产物。
2	两条带	在同一个样品中可以看到两个条带，强度基本相等。
3	多条带	在同一个样品中可以看到三个或者三个以上的条带，强度基本相等。
4	模糊不清	PCR 失败，未能得到明确的产物，或者是样品中的化学污染物使得样品不能进行正常电泳。这种情况几乎不会出现。
5	弱的条带	与 DNA 标记物相比，PCR 产物的强度不高。
6	冗余条带	PCR 产物含有两条或更多条带，但是可以观察到一条主要的条带，其他冗余的条带有可能干扰阵列的杂交。

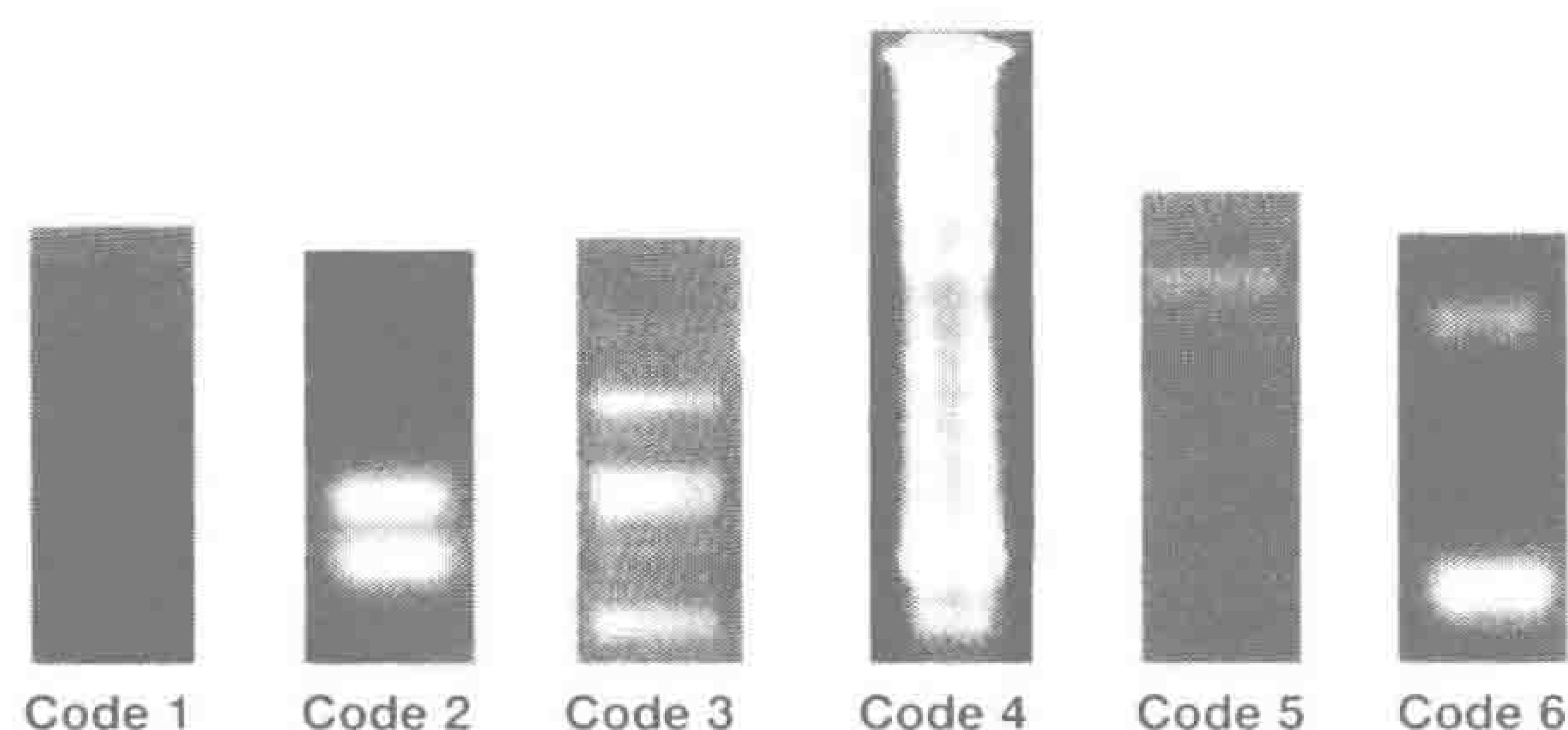


图 1-3　PCR 产物扩增过中出现的问题示例

将图像转换成数值代码，就可能把该信息加入到克隆注释工作表中，从而应用于对微阵列实验结果的评价

尽管一些 PCR 产物会出现多种缺陷，但每一样品只能有一个分值代码。两条带（2）和多条带（3）优先于弱的条带（5）。例如，当一个样品中出现两条具有相同强度的弱的条带时，样品标记为双条，而不作为弱的条带（5）样品。

26

标准	分析中使用的 DNA 标准溶液/μl	每个孔中 DNA 的总量/ng
1	100.00	200.00
2	100.00	100.00
3	100.00	50.00
4	100.00	25.00
5	100.00	12.50
6	100.00	6.25
7	100.00	3.12
8	100.00	0.00

27

附加方案：用 PicoGreen 进行定量

多种方法可用于对扩增产物的定量：

- 通过把琼脂糖凝胶中条带的强度与标准样品进行比较来估计。
- 测量在 260 nm 处的光吸收。该方法的缺点是灵敏度相对较低（为了得到在线性范围内的结果，需要 10%～20%的 PCR 产物），核苷酸及单链核酸对信号有影响，且有来自污染物质的干扰。后两个问题可以通过扩增后对产物进行适当的纯化而减轻。
- 使用高灵敏度的染料，如 PicoGreen，可以在 pg/ml 范围内对双链 DNA 进行检测，测量的线性范围为 25 pg/ml 至 1000 ng/ml。

如果始终可以得到类似于图 1-2 所示的结果，那么只要检测少量纯化的样品在 260 nm 处的光吸收，就可以估算出整批反应的平均产率。然而，在产量可变时，最好用 PicoGreen 测量全部样品的产率，并对样品的体积作相应调整，这样更为准确。产品间的变化使得这个过程很繁琐，但这对生产高品质的微阵列来说是非常必要的。即使对样品的体积不做调整，了解哪些样品来自于微弱的或者失败的扩增反应，对解释微阵列实验结果也是非常有用的。

附加材料

铝箔

DNA 标准

制备 10 ng/μl DNA 标准的储存液。把样品按 100 μl 分装后放在－20℃下冷冻保存。

PicoGreen 双链 DNA 定量试剂盒（P-7589，Molecular Probes）

多孔板（黑色 96 孔，带透明底部）（25096，E&K Scientific）

3×SSC，用 20×SSC 稀释配制（15557-036，GIBCO）

荧光分光光度计微量滴定板读板器（Tecan Spectra Fluor）

TE（pH 7.5）

方法

1. 由储存液配制两份 DNA 标准液。

2. 从每个 PCR 产物或纯化产物中取 1 μl，转移至新的 96 孔板的孔中，每一孔中含有 50 μl 的水或 TE（pH 7.5）。

如果样品不在当天使用，则要用铝箔带盖好多孔板，并于－20℃下冷冻保存。在进行后续分析时不要丢弃这些临时性的多孔板，可能需要对它们再次检查。

3. 从“临时性多孔板”移取 5 μl 每种 DNA 样品，放入一块黑色的 96 孔板中，并向每个样品孔中加入 95 μl 的 TE（pH 7.5），使溶液总量为 100 μl。
4. 按照 1∶200 用 TE（pH 7.5）稀释 PicoGreen 染料储存液，配制成 PicoGreen 工作液。
5. 向每个样品孔中加入 100 μl 的 PicoGreen 工作液，并把多孔板在多孔板摇床上振动 5 min 以充分混合（振动之后要离心 1 min，以便回收黏在孔壁上的液体），也可以利用多孔板读板器自身的振动功能，在检测之前振动多孔板。
6. 用铝箔盖住样品板，防止样品受光照。
7. 用多孔板读板器（如 Tecan Spectra Fluor）进行样品检测，并将数据输出至 Excel 或 SigmaPlot 软件中。
8. 建立标准曲线，并计算 DNA 溶液的浓度。

方案 5

28

点样法阵列备用 DNA 的重新悬浮及储存

Andrew Holloway，Maria Murphy（Peter MacCallum 癌症研究所，东墨尔本，澳大利亚）；
Settara Chandrasekharappa（国立人类基因组研究所，国立卫生院，贝塞斯达，马里兰州，20892）

点样开始前，必须将可以立即使用的阵列材料再次悬浮。对重新悬浮 DNA 溶剂的选择取决于点样针或传送装置的形状、所使用的点样基底、点样时样品板暴露时间的长短以及点样设备所具备的环境控制类型和级别。由于点样结果受诸多因素的影响，因而最好的方法是通过实验来确定特定条件下适合的溶剂和点样条件。

点样溶液

最常用的 DNA 点样溶液有 3×SSC 或 150 mmol/L 的磷酸缓冲液（pH 8.0）。这两种点样液均可按下述方法进行补给：

DMSO（最高为 50％ *V/V*）可以用来减少点样中的蒸发，从而改善样点的形态，有利于 DNA 的变性（Hegde et al. 2000）。可以加入甜菜碱（1.5 mol/L）来提高 DNA 的结合效率，增加样点的均匀性，同时降低背景（Hegde et al. 2000）。

用装备了针和环系统的机器点样时，十二烷基肌氨酸钠或十二烷基硫酸钠（SDS）（0.05％～0.1％）可用来降低表面张力，促进样点扩散。这样可以减少生成圆环形（doughnut）样点，得到更加均匀的样点形态。十二烷基肌氨酸钠也可以与中空的点样针配合使用。例如，ArrayIt 推荐应用 ArrayIt Stealth Microspotting 点样针时使用 0.005％的十二烷基肌氨酸钠。

体积

加样板（printing plate）板孔中溶液的体积很关键。板孔中溶液的量要能保证每次点样针中溶液填充时能够保持一致，同时要考虑到在点样过程中溶液的蒸发。反之，板孔中溶液过多又会引起点样针外周过分润湿，导致样点形态差，发生滴液现象。384 孔加样板中，板孔内溶液体积一般为 5 μl 或 6 μl（不超过 1 mm 深）。

阵列材料的浓度

相对于靶分子，探针 DNA 要过量，以确保得到良好的杂交信号和定量时最大的动态范围。然而，过量的浓缩溶液会造成昂贵材料的浪费，并导致弥散或拖尾现象发生（见第 3 章）。点样液中 DNA 的浓度通常为 100～250 μg/ml。如果 DNA 浓度低于 100 μg/ml，灵敏度会下降，同时会低估样品间差异性表达的程度（Yue et al. 2001）。

29 **点样过程中的保存**

由于在点样过程中以及两次点样操作之间，点样液的蒸发是变化的，因而每次点样结束后，最好在真空浓缩器中将点样材料彻底干燥（或者在 37℃的孵育箱中过夜）。在 −20℃下，将干燥后的材料密封于多孔板中保存。点样前加入蒸馏水使材料复原（reconstitute），要考虑到上次点样时所用掉的样品体积。一般情况下，中空点样针［如 ArrayIt（Stealth Microspotting Pins，SMP3）］每次点样消耗的体积为 0.25 μl，Affymetrix 的针和环式点样仪每次消耗液体体积为 1 μl。在多次点样、干燥、补充和重新悬浮的循环操作中，可以使用同一块加样板。然而必须注意到，残留的材料不能随时间变化而降解或交叉污染，以避免降低所制备的 DNA 微阵列的质量。

本实验方案讲述了纯化 PCR 产物（方案 1～方案 3）的重新水化。

材料

缓冲液和溶液

储存液、缓冲液和试剂的成分见附录 2。

将储存液稀释至合适浓度。

选择的点样缓冲液［3×SSC 或 150 mmol/L 的磷酸缓冲液（pH 8.0）］

点样液也可以从一些公司购得（如 Telechem 公司产品，www.arrayit.com）。

核酸和寡核苷酸

来自扩增反应纯化后的样品（方案 1～方案 3）。

专用设备

用于密封点样板的黏性胶带（如 7742000，ICN）

带有适配器的台式/医用离心机

这些离心机可以离心深的和标准的 96 孔微量滴定板（如 Eppendorf 5810R，来自 Beckman Coulter 的 Allegra 6R，其他同类产品）。

方法

1. 向带有纯化 DNA 探针的微孔板中加入适量体积的点样缓冲液。用黏性密封胶带覆盖多孔板，并在 1300 *g* 下离心 2 min。

   关于点样缓冲液加入的量和浓度，请参考本方案的介绍部分。

2. 把样品在微量板振荡器上混合 5 min。
3. 在 1300 *g* 下离心 2 min，回收黏在孔壁上的 DNA 样品。
4. 将样品于 4℃下静置过夜。
5. 重复步骤 2 和步骤 3。

方案 6

30

从酵母中分离和扩增用于阵列的材料

Vishwanath Iyer（得克萨斯大学，奥斯汀，得克萨斯）

酵母基因组 DNA 是通过 PCR 扩增编码区和其他 *S. cerevisiae* 位点的模板。第一个测序的酵母基因组是 S288c 菌株（*MATα gal2 mal*）。该菌株的 DNA 通常被用作扩增模板的来源，而且以此序列来设计寡核苷酸引物对。关于进一步的信息，请见关于**酵母探针**的信息栏。本实验方案分为三个阶段：

- 酵母基因组 DNA 的分离
- 基因组 DNA 的纯化和分析
- 通过 PCR 扩增基因组 DNA

为了使单次 PCR 循环扩增的产量达到最大，建议对于每个基因或微阵列单元，PCR 扩增的体积为 100 μl；每个反应需要的酵母基因组 DNA 模板的量为 40 ng，因此扩增全酵母基因序列需要约 250 μg 模板 DNA。本方案所述的方法可以从 500 ml 在丰富培养基中生长接近饱和的酵母培养物中提取约 1 mg 基因组 DNA。基因组 DNA 也可以用商品化的试剂盒进行分离（如来自 Qiagen 的试剂盒）。此外，S288c 基因组 DNA 可以从一些公司购买得到，如 Research Genetics（40802）和 Novagen（69240-3）。

对于酵母基因组成分扩增，基因组引物集可以以 96 孔微量滴定板的形式得到，并且所有的操作处理均按照这一方式进行（表 1-3）。能否在合理的时间内完成如此大量的 PCR 扩增，取决于是否有可供高通量操作的合适的基本设施。在着手操作之前，建议用一小部分（如 4 块）96 孔引物板对整个程序进行预实验，以确保每一步都能成功。可以提前给 PCR 和沉淀用的板标上记号，以便简化过程。

由于扩增反应的反复、引物合成或者 PCR 的问题，可能会出现个别失误。使用这里介绍的扩增引物集，扩增反应成功率应该超过 95%。

表 1-3 用于扩增酵母基因组片段的引物集

种 类	数 量	可用次数	商品目录编号
板载 ORF Genepairs	10 μmol/L 混合液 30 μl	3	40603. 1P
管载 ORF Genepairs	20 μmol/L 储存液 200 μl	50	40603. 1T
板载 Intergenic Genepairs	10 μmol/L 混合液 30 μl	3	40903. 1P
管载 Intergenic Genepairs	20 μmol/L 储存液 200 μl	50	40903. 1T
板载 pORF Genepairs	10 μmol/L 混合液 50 μl	5	40510. 1P
管载 pORF Genepairs	20 μmol/L 储存液 200 μl	50	40510. 1T

引物集可以从 Research Genetics 获得。

材料

注意：标记有〈!〉的材料的正确处理方法请参见附录 3。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

β-巯基乙醇（0. 28 mol/L）〈!〉

乙醇（95%和 70%）

据报道，有些批次的乙醇（100%或 200 度）中含有会使微阵列试验中产生很深荧光背景的污染物质，95%的乙醇不存在这个问题。为避免这个问题，可以用 95%的乙醇或者用已证明不会产生高荧光背景的无水乙醇。Warner-Graham 公司（Cockeysville，Maryland）出售的无水乙醇（USP 级乙醇，64-17-5）效果很好。

裂解缓冲液（100 ml）

0. 9 mol/L 山梨醇

0. 1 mol/L Tris-Cl（pH 8. 0）

0. 1 mol/L EDTA

$MgCl_2$（25 mmol/L）

醋酸钾（5 mol/L）

重悬缓冲液 Ⅰ（50 ml）

50 mmol/L Tris-Cl（pH 8. 0）

20 mmol/L EDTA

重悬缓冲液 Ⅱ

50 mmol/L Tris-Cl（pH 8. 0）

1 mmol/L EDTA

SDS（10%）〈!〉

醋酸钠（3 mol/L，pH 5. 2）

酶和缓冲液

10×扩增缓冲液（可从 Applied Biosystems 得到，N808-0190）

此外，按照如下配方配制的 10×扩增缓冲液效果很好：

500 mmol/L KCl

100 mmol/L Tris-HCl (pH 8.3)

在步骤 20 中单独加入 $MgCl_2$。缓冲液分装后保存于－20℃。

RNaseA (10 mg/ml) 分装后保存于－20℃。

Taq DNA 聚合酶 (5 单位/μl)（如来自 Applied Biosystems 的 Amplitaq Gold）

Zymolyase（用裂解缓冲液配制 0.5 mg/ml 溶液）

使用前新鲜配制。

核酸和寡核苷酸

dNTP（四种 dNTP 的 100 mmol/L 溶液，27-2035-03 Amersham Pharmacia）

在 96 孔板（引物板）上的正向和反向 PCR 引物对

所提供的引物在 TE 中，浓度 20 μmol/L，保存于－20℃备用。

选择合适引物对的信息，参见关于**酵母探针**的信息栏和表 1-3。

细胞

32

Saccharomyces cerevisiae S288c 菌株

该菌株可以从 ATCC（26108）或 Research Genetics（95500）购得。

培养基

YPD 培养基（制备 500 ml）

专用设备

带滤芯吸头

在所有未使用多道移液器的移液步骤中使用带滤芯的吸头，以避免先前残留物的污染。

带适配器的台式/医用离心机，能离心深的和标准尺寸的 96 孔微量滴定板（如 Eppendorf 5810R，来自 Beckman Coulter 的 Allegra 6R 或同类型的离心机）

PCR 板（96 孔）和适用于加热盖的密封垫（如 MJ Research 的 Multiplate 和 Microseal A）

吸头（一次性的，放在 8×12 的吸头架上以适合于移液器）和移液槽

移液器（容积可调，12 通道，2.5 μl 和 200 μl 容量）

移液器（单通道，可调式）和带滤芯吸头

热循环仪（带加热盖的 96 孔板）

来自 MJ Research 的 Tetrad 可以容纳 4×96 孔板，容许实现高通量。

另一个选择是来自 Applied Biosystems 的 GeneAmp 9700。

水浴锅（设置在 65℃）

附加试剂

本实验方案需要附加的试剂和设备，用于纯化和分析方案 3 和方案 4 中所列的 PCR 产物。

方法

酵母基因组 DNA 的分离

1. 用 1 ml 或 2 ml 过夜培养的 *S. cerevisiae* S288c 菌株接种 500 ml YPD 培养基。将培养物置于振荡的孵育箱中，30℃过夜生长。生长结束时，培养物应该接近饱和（OD_{600}约为 5.0）。
2. 室温下 2000 *g* 离心 10 min，收集细胞，弃去上清液，用 25 ml 裂解液重悬细胞，用 10 ml 的血清移液管反复吸取和（或）涡旋振荡使细胞完全重新悬浮。
3. 向悬浮的细胞中加入 2.5 ml 浓度为 0.28 mol/L 的β-巯基乙醇和 2.5 ml 新鲜配制的 Zymolyase（0.5 mg/ml，用裂解缓冲液配制），涡旋振荡使溶液混匀。
4. 把悬液分成等份置于两个小的离心管（30 ml）中。37℃下，以 200 r/min 的速度将离心管振荡孵育 1 h。

   该步骤消化细胞壁得到原生质体。
5. 室温下，2000 *g* 离心 5 min，收集原生质体，弃去上清液。
6. 在每支离心管中，将原生质体再次悬浮于 10 ml 重悬缓冲液Ⅰ，反复吸取和（或）漩涡振荡以确保完全重新悬浮。

33

7. 向每支离心管加入 1 ml 10%的 SDS，振荡混匀，将离心管在 65℃水浴中孵育 30 min。

   本步骤裂解原生质体。
8. 向每支离心管加入 4 ml 5 mol/L 的醋酸钾，倒转离心管数次以混匀溶液。将离心管置于冰上 30 min 至 1 h。
9. 4℃下以 5000 *g* 离心 15 min，去除沉淀的 SDS、蛋白质以及细胞碎片，小心收集上清液，避免搅动沉淀。把上清液（总共约 30 ml）平均分到 4 支离心管中。
10. 向每支离心管加入两倍体积（约 15 ml）的 95%乙醇，核酸应立即沉淀出来。将离心管在室温下放置 5 min。
11. 将离心管在 4℃下以 5000 *g* 离心 30 min，使 DNA 呈小团块状。用 1 ml 重悬缓冲液Ⅰ将各个小团块再次悬浮。将上清液合并至一支离心管内（共约 4 ml）。重悬时需要用 1 ml 可丢弃的移液器吸头破碎小球，并反复吸打。

基因组 DNA 的纯化和分析

12. 加入 20 μl RNase A（10 mg/ml），将离心管在 37℃下孵育 30 min。
13. 加入 400 μl 3 mol/L 的醋酸钠（pH 5.2）和 11 ml 95%的乙醇。将离心管在室温下放置 5 min。此处（或在第 10 步中）沉淀出来的 DNA 也可以在－20℃下过夜放置。
14. 将离心管在 4℃下以 5000 *g* 离心 30 min，使 DNA 呈小团状。
15. 用等体积的 70%乙醇洗涤小团。真空抽吸除去乙醇溶液，并将开口的离心管直立在工作台上至 DNA 小团半干。用 4 ml 重悬缓冲液Ⅱ使 DNA 小团重悬。

    DNA 完全溶解可能需要数小时。将离心管竖立在温度为 60℃的水浴中可以加速这一

过程。

16. 用分光光度法定量 DNA 并将它按最终浓度为 400 ng/μl 重悬。
17. 用琼脂糖凝胶电泳分析少量制备的 DNA，以确保经过溴化乙锭染色后观察到的 DNA 条带强度与计算的浓度值一致。典型的凝胶图像如右图所示。



用 PCR 扩增基因组 DNA

18. 将含有引物的多孔板从冰箱中取出，于室温下或盛有温水（约 45℃）的浅碟中解冻。待引物融化后，将引物板以 2000 *g* 离心 2 min，使溶液汇聚在底部。
19. 仔细剥下引物板上的覆盖物，用多道移液器上下吸打两三次，充分混匀储藏板（stock plate）上的引物。移取 5 μl 每种正向引物和反向引物到定位与引物板一致的 96 孔 PCR 板上。

    每次操作母板（source plate）时要更换一组新的吸头。

34

20. 对于每块 96 孔板，在一次性的多通道移液器储液皿中按顺序加入下列物质，作为扩增反应混合物。该反应混合物足够供给 4 块 96 孔板使用，也可以在一个储液皿中混合。

水	7.05 ml
10×PCR 缓冲液	1 ml
25 mmol/L $MgCl_2$	0.8 ml
100 mmol/L 四种 dNTP 各自的储存液	每种 25 μl
酵母基因组 DNA（400 ng/μl 储存液）（来自步骤 16）	10 μl

    扩增全基因组序列时，较为方便的做法是配制大量包含上述成分的主要混合物（master mix）储存液，分装成每份 36 ml，在 -20℃下冷冻保存。每份足够供给 4 块板使用。密封盖旋松的试管用微波处理 20 s，以融化冰冻的液体。将融化的混合物转移到多通道移液器储液皿中，在即将把混合物加入到引物中之前，首先加入 *Taq* 聚合酶。

21. 向每块 96 孔板反应混合物中加入 40 μl *Taq* 聚合酶（5 单位/μl）。沿着储液皿的宽度方向均匀地加入黏性的酶。用移液器上下吸打数次以充分混匀。
22. 用多通道移液器转移 90 μl 含有 *Taq* 聚合酶的反应混合物到带有引物的 96 孔 PCR 板上的每一行中。加入反应混合物时，用移液器上下吸打两三次，充分混匀最终溶液。加入每种反应组分时都要更换一套新的吸头。
23. 用密封垫盖住 PCR 板，放入热循环仪器中。
24. 按照下面的变性、复性与延伸时间和温度进行扩增反应：

循环次数	变　性	复　性	延　伸
1	95℃，30 s		
36	94℃，30 s	56℃，45 s	72℃，3 min 30 s
最后一个循环		72℃，5 min	

这些条件是为使用 Research Genetics 的引物扩增酵母基因序列而进行优化的，适合在本实验方案中介绍的专用设备和反应；然而，这些条件在用于其他类型的设备和反应时可能需要进行调整。在开始时，增加额外的 94℃变性步骤，以及 72℃和 25℃下的最终延伸，不会明显影响产量和成功率。

关于扩增产物的分析、定量和纯化，见方案 3 和方案 4。

35

信息栏：克隆集

表达序列标签

表达序列标签的来源

传统的基因发现方法，如隐性或显性变异的作图和克隆，或者根据生物化学上确定的蛋白质分离基因，已经定义了人和模式生物体中表达序列的 10%～20%。截至 2001 年，大约有 7000 个人类基因被定义为“已命名的”，并且有与其相关的附加生物学信息。然而，通过应用大规模的 cDNA 测序来产生表达序列标签（EST），使得基因发现的速度大大地加快了。最初由 Craig Venter 和他的同事们所倡导（Adams et al. 1991），目前，企业和大学实验室中都在进行对大量物种的 EST 鉴别和分类的研究项目。到目前为止，大部分精力集中在人类的 EST。1991 年，通过 Merck 公司和 Washington 大学的合作，向 GenBank 中添加了 480 000 个人类 EST 序列。到 2001 年，通过 Washington 大学和 IMAGE 联盟协定［（Lennon et al. 1996），http://image. llnl. gov］成员及其他项目如癌症基因组解剖计划（Cancer Genome Anatomy Project，CGAP，http://cgap. nci. nih. gov）的努力，该数字已经增加到至3 000 000。

EST 命名法

关于 EST 克隆，有许多混淆的名称和标识符。表 1-4 有助于弄清楚这些标识符[承蒙美国种质保藏中心（ATCC）允许]。值得注意的是，以 HG、HT 或 THC 开头的名称并非克隆的唯一标识符，它们分别表示由拼接 EST 序列得到的人类基因、人类转录子或假定的人类一致性序列的组合标志符。下列是特定的标识符：ATCC 号，MGC 号，克隆名称，GenBank 或 dbEST 标识符，GDB 探针标识符。

表 1-4　与 EST 克隆有关的名称和其他标识符

标识符	格　式	与克隆对应关系	示　例
ATCC 号	5～7 位数	唯一	105303
MGC 号	4～7 位数	唯一	20317
克隆名称或别名	包括文字与数字	每个克隆可能不止一个	NIB345
IMAGE 克隆 ID	5～7 位数	每个克隆可能不止一个	60800
GenBank 登记号	大写字母后接 5 位数字或两个字母后接 6 位数字	每个克隆可能不止一个	T12121，AA234567
EST 号	EST 后接数字	每个克隆 0 个或数个	EST149990
Washington 大学序列 ID	小写文字与数字串后接句点，然后 r 或 s，最后一位数字	每个克隆 0～2 个	Yb34g01. s1
Unigene	两个字母后接 1～6 位数字	唯一	Hs. 1254

EST 数据库作为定义新基因的工具

从大规模测序项目得到的 EST 序列存放在 GenBank 的一个分支 dbEST（http://www.ncbi.nlm.nih.gov/dbEST/index.html）中，有一个名为 UniGene 的自动处理用来比较 EST 和以与鸟枪法测序类似的方式将重叠序列合并到各自的类别当中去（http://www.ncbi.nlm.nih.gov/UniGene/index.html）。每次这一过程重复执行时（称之为一次 UniGene build，http://www.ncbi.nlm.nih.gov/UniGene/build.html），一些附加的序列可用于把先前的非重叠类别合并到单个基因中，从而取消一些类别，并进一步改进对所表达基因的定义。从基因组测序得到的信息可被用来进一步减少类别的数量。相反的，简缩后的类别可以为基因组序列注释提供信息，并使之变得容易进行。在理想情况下，每个类别会对应于一个基因，但是有许多个非重叠的类别也可能代表了大型的基因、表达水平低的基因，或者是基因的可变剪接的转录子。不仅如此，序列比对程序的差错会产生虚假的类别（过度聚类）。在可以预期的将来，由序列得到的类别数目可能会超过产生这些序列的各个基因的总数。IMAGEne 数据库（http://image.llnl.gov/image/imagene/current/bin/search）将人类的 IMAGE 克隆与 GenBank 中的非冗余记录比对，并获得一组相关的 IMAGE 克隆，按全长、插入片段大小及文库质量排列。一个 Java 程序使得用户可以查看实际的序列比对。

其他一些数据库汇集了形成一个基因或 mRNA 转录子的相邻区的一致性序列。这些通常可以根据基因名称或者 GenBank 登记号来查询。TIGR 数据库（http://www.tigr.org/tdb/index.shtml）将人类的 EST 合并到人类转录子（HT）中，用 EGAD（expressed gene anatomy database，http://www.tigr.org/tdb/egad/egad/shtml）中的基因名称和基于 HGI（human gene index，http:www.tigr.org/tdb/hgi/index.shtml）中 EST 的相邻序列数据的 TIGR 人类叠联群（THC 集合）来描述。THC 集合非常有价值，因为它们表示了公认的基因序列。通过连接两种数据库的集成使得可以应用基因名称、细胞内作用角色、序列数据、EST 标识符，或者 GenBank 登记号来进行检索。HGI 包含了许多不同来源的克隆，并且显示了克隆对应于人类转录子的相对位置。EGAD HT描述符提供了用转录子编码序列定义的核苷酸的范围。

建立 The allgenes 数据库（www.allgenes.org）的目的，是利用公开的数据提供一个综合的、关于所有已知和预测的人类和老鼠基因的数据库。它试图把各种类型的数据，包括 EST 序列、基因组序列、表达数据，以及功能信息集中在一起。Allgenes 为预测的人类和鼠基因提供了一个详细的索引。当前，基因的预测来自于转录子，通过对 EST 和 mRNA 序列的聚类和拼接来预测。这些 EST 和 mRNA 类别是来自 DoTS（转录序列数据库，database of transcribed sequences）最新公布的数据，该数据库由宾夕法尼亚大学的计算生物学和信息学实验室（http://cbil.upenn.edu/）建立。stackdb（www.sanbi.ac.za/Dbases.html）是一个相似的尝试，旨在利用公开的序列数据建立一个人类转录子的完整的数据库（Miller et al. 1999；Christoffels et al. 2001），提供关于转录子表达的局部分布信息。它是由 SANBI（the South African National Bioinformatics Institute）和 NCGR（National Centre for Genome Resources）建立的。

MGC（mammalian gene collection，http://mgc.nci.nih.gov）为人和鼠的表达基

因提供了一套完整的全长（可读框或 ORF）序列和 cDNA 克隆。MGC 是由 NIH 资助的项目，它支持 cDNA 文库、克隆及序列的制备。NIH MGC 试图为每个人和鼠基因辨别和测序一条有代表性的全长 ORF 克隆。自从 2000 年以来，MGC 建立了超过 80 个 cDNA 文库，富集了从人组织和细胞系，及小鼠组织中获取的全长 cDNA。每个库的质量首先用有限数目克隆的 5′和 3′EST 分析来评价。另外 10 000～30 000 个 5′EST 的阅读由达到该项目质量标准的每一个文库来产生。有多种算法可以用来选择公认的全长 ORF 克隆。

37

到 2001 年 2 月，生成了超过 500 000 个 5′EST 序列的阅读，并且有一套非丰余的、数目超过20 000的人类和小鼠公认的全长 ORF 克隆被汇集到 MGC 中。这些克隆的插入片段已经被全部高准确度地测序，并且评价了 EST 是否覆盖了一个完整的可读框架。在本书编写期间，有超过 3000 个克隆完成全部测序，选择的克隆约有 75%经证实拥有全长 ORF。MGC 的所有资源，包括 EST 序列、全长插入序列和克隆，全部可以被公众获取。全长克隆对微阵列实验的价值可能更多地在于其序列信息和把它们应用于功能验证，而不是作为探针的资源。全长克隆很难有效地大量扩增，并且这类探针比局限于 3′UTR［非转录区（untranslated region)］序列的克隆更加容易与相关联序列发生交叉杂交。

其他生物，如鼠（http://genome. wustl. edu/est/mouse _ esthmpg. html)、*Drosophila melanogaster*（http://www. fruitfly. org/sequence/genomic-clones. html）以及 *Arabidopsis*（http://www. cbc. umn. edu/ResearchProjects/Arabidopsis/index. html）EST 的测序，就绝对数量来说，要落后于人类。因此这些物种中的聚类受到更多的限制，并且相应克隆集的质量也因此更加受限。有趣的是，对大量 *Drosophila* 睾丸cDNA 的测序鉴别出许多与基因组 DNA 序列相关的克隆，但先前却没有被预测为表达的基因（Andrews et al. 2000)。这就突出表明了当前用于分析复杂基因组的算法的脆弱性。

EST 用作微阵列的资源

由每个聚类类别的单一代表组成的克隆集，通过经过许可的供应商向研究人员分发（http://image. llnl. gov/image/html/idistributors. shtml)，其中最全面的克隆集是由 Research Genetics 和 Resource Center of the German Human Genome Project（RZPD）来分发的（表 1-5)。通过克隆纯化和重新测序做了大量的努力，以减少克隆集的错误率。LION Bioscience（http://www. lionbioscience. com）建立了数种专门为微阵列分析而优化的小鼠、大鼠以及狗的克隆集：它们相互之间大小相似（通常为 200～600 bp)，对应于 mRNAs 的 3′UTR，并且不含有重复元件和 poly(A)区［poly（A）tracts］（表 1-5)。这些特性将使得克隆与克隆之间在标记强度上的差异减小，并将交叉杂交问题减至最小。Incyte Genomics 过去曾经是一个主要的有效克隆集的提供者，在 2001 年 5 月停止了克隆集的供应，但还继续供应来自这些克隆集的，用于微阵列点样的现成材料。除了 IMAGE 联盟协定外，其他主要的 cDNA 克隆集贡献者包括RIKEN（http://genome. rtc. riken. go. jp/home. html)、TIGR（http://www. tigr. org）以及其他的独立单位（http://www. atcc. org/SearchCatalogs/tasc2. cfm＃who)。在大多数此类情况下，单个克隆或者克隆集都对公众开放，虽然当中有一些使用时需要经过授权或共享由此生成的任何知识产权。

表 1-5 克隆集

38

	Research Genetics/Invitrogen[a]	Resource Center of the German HGP[b]	LION Biosciences[c]
目前可得到的	**已经核实的人类序列** 已知的基因收藏约 7500 EST 基因聚类收藏约 40 000 该收藏包括超过 45 000 个已核实序列的克隆，来自IMAGE联盟协定/LLNL Human 收藏，包括 NCI CGAP 收藏。每个 cDNA 克隆选自基于 3′端序列优先的一个 UniGene 聚类类别和一个 NCI-CGAP 克隆。利用 BLAST 分析把序列与 NCBI UniGene 数据库相匹配来确认一致性	**人类克隆集** **RZPD Human Unigene1**：36 425 个克隆（冗余度>1.8 倍，基于 Unigene 聚类） 可供索取 http://www.rzpd.de/cgi-bin/db/showLib.pl.cgi/response?libN0=950 **RZPD Human Unigene2**：大于 72 000 个克隆（不完全清楚其冗余度） Human Unigene2 克隆被用作表达谱研究（插入片段 400～1500 bp，平均为 800 bp；基因的位置强烈偏向于 3′端；只用一对引物扩增全部克隆） 基于 Unigene 聚类。RZPD Human Unigene2 已经全部再测序（2001 年底完成）以核实和去除冗余部分 花费：105 000 欧元 **RZPD Human Unigene3**：48 000～53 000 个克隆 Human Unigene3 是 Human Unigene2 的一个已核实序列的子集，没有任何冗余序列，为表达谱研究而优化（见 human Unigene2） 基于 Unigene 聚类并作重要修改，2002 年初可以获得	**ArrayTAG Mouse A**（状况：2001 年 7 月） 已核实序列的 cDNA 克隆：10 000个 7198 个与 NCBI Unigene 聚类相匹配，其中： 2207 个描述中显示“EST”； 141 个被称作“DNA 片段”； 2354 个的描述类似“Riken cDNA…”； 2496 个有明确的注释； 2800 个阵列 TAG 代表了尚未知的转录物。插入范围 200～600 bp（平均值 303 bp，标准差 83 bp）。 提供全部插入序列
	已经核实的小鼠序列 基因类别收藏约 16 000 该收藏包括超过 15 000 个已核实序列的克隆，来自IMAGE联盟协定/LLNL Mouse 收藏，包括 BMAP 收藏。序列的核实过程与上述对人类克隆的核实相同	**小鼠克隆集** **RZPD Mouse Unigene1**：24 500 个克隆（冗余度>2 倍） 可使用两种不同的引物对扩增克隆，平均插入片段约 680 bp 基于 DKFZ 理论生物信息学系进行的聚类研究 http://www.rzpd.de/cgi-bin/db/showLib.pl.cgi/response?libNo=952 **RZPD Mouse Unigene2**：大于 50 000 个克隆（其冗余度不清楚） 基于 DKFZ（Deutsches Krebsforschungzentrum）理论生物信息学系进行的聚类研究；为表达谱研究而优化（见 human Unigene2）；2001 年底可以获得（译者注，现已升级为 Mouse Unigene Set RZPD2.1） **RZPD Mouse Unigene3**：大于 40 000 个克隆（全部得到序列核实，没有冗余序列） Mouse Unigene3 是 Mouse Unigene2 的一个子集，基于 Unigene 聚类并作重要修改；为表达谱研究而优化（见 human Unigene2）；2002 年初开始供应	**ArrayTAG Rat A**（状况：2001 年 7 月） 已核实序列的 cDNA 克隆：10 000个 7027 个与 NCBI Unigene 聚类相匹配，其中： 3780 个描述中显示“EST” 500 个被称作“EST 高度相似” 1025 个归类为“EST 中等”或“EST 弱相似” 1722 个有明确的注释 2900 个阵列 TAG 代表了尚未知的转录物 插入范围 200～600 bp（平均值 322 bp，标准差 60 bp）。 提供全部插入序列

39

续表

	Research Genetics/Invitrogen[a]	Resource Center of the German HGP[b]	LION Biosciences[c]
目前可得到的	**已经核实的大鼠序列** 基因聚类收藏：约 16 000 当前，来自于 Iowa 大学（Dr. Bento Soares）的收藏包括超过 15 000 个已核实序列的褐家鼠（*Rattus norvegicus*）克隆，由用于 EST 测序的标准化和减数大鼠（Sprague-Dawley）cDNA文库组成。对由此得到的克隆进行聚类，用于产生一个唯一的大鼠基因收藏，来自成熟的胚胎组织	**大鼠克隆集** **RZPD Rat Unigene1**： 28 896 个克隆（冗余度很可能小于 2.5 倍） 基于 Iowa 大学进行的聚类研究；目前已有供应，请向厂商询价 http://www.rzpd.de/cgi-bin/db/showLib.pl.cgi/response?libNo=953	**ArrayTAG Dog A** 已经核实序列的 cDNA 克隆：10 000 个 犬 cDNA 的注释策略是基于对人的 NCBI-Unigene 聚类的同源性搜索和由 LION 的序列注释包 bioSCOUT 产生的注释数据。根据给定的严格度标准，约 15% 的狗 arrayTAG 序列与它们对应的人 NCBI-Unigene 聚类相匹配，约 35% 由自动的 bioSCOUT 执行赋予明确的注释 插入片段的范围 400～900 bp（平均值 602 bp，标准差 60 bp）
	果蝇（*Drosophila melanogaster*） EST 收藏：约 55 000 *D. melanogaster* EST 由伯克利果蝇基因组计划（BDGP，Berkeley Drosophila Genome Project）构建。克隆是由各种组织产生，利用两个不同的载体。关于详细信息，请参见 http://www.fruitfly.org	**拟南芥（*Arabidopsis*）克隆集** **RZPD *Arabidopsis* Unigene1**：小于 15 000 个克隆（没有冗余序列） 基于 RZPD 聚类研究；2002 年第二季度开始供应	—
	酿酒酵母（*Saccharomyces cerevisiae*） ORF 专用引物：约 6000 酿酒酵母 ORF 专用引物被设计用于从基因组 DNA 中扩增包括起始密码子和终止密码子的整个编码区域。所有正向引物的 5′ 端都被设计以 5′-GGAATTCCAGCTGACCACC 序列开始，随后是起始密码子 ATG 和 ORF 的最初大约 25 个碱基。所有反向引物的 5′ 端都被设计以 5′-GATCCCCGGGAATTGCCATG 序列开始，随后是野生型终止密码子和 ORF 的最后大约 25 个碱基的反向互补序列	**各物种已命名的（已知的）基因** 从所有物种的克隆中选取出的子集，以产生可以代表来自各个物种全部已知基因的克隆集；2002 年初开始供应	—
	爪蟾（*Xenopus laevis*） 有多种 *Xenopus laevis* 的收藏可供利用。关于详细信息，请参见 www.resgen.com	**定制的克隆集** 根据用户递交的存取号和（或）序列来提供克隆。递交的克隆是上述的 Unigene 集的子集，因此序列已经核实	—

续表

	Research Genetics/Invitrogen[a]	Resource Center of the German HGP[b]	LION Biosciences[c]
目前可得到的	**TIGR EST 克隆** 从 TIGR EST 收藏可以获得许多克隆，包括病毒、真细菌（病原体和非病原体）、古菌（即所谓第三界生命）、真核生物（植物、动物、真菌和原生生物如疟疾寄生物等）。更多的信息，请参见 www. tigr. org	—	—
T1 污染	已经核实的序列克隆转化到抗 T1 的宿主并放入已经核实序列的克隆收藏中	对所有克隆都进行了噬菌体侵染和非 *E. coli*污染检验，只有在通过各种检验之后才能提供	在发货之前对全部复制的收藏随机取样，检验噬菌体 T1、T4 和 λ 的污染。ArrayTAG 克隆集没有噬菌体污染史
提供形式	所有收藏以 96 孔板的形式提供	克隆集可以多种不同的形式和格式获得：滤纸上的高密度克隆阵列、DNA 阵列、PCR 产物集和以干冰冻存的 384 孔格式的活菌株	以甘油储存液形式低温冷藏于 96 孔板中
提供的信息	针对每个克隆集，提供克隆标识符、基因名称（如果已知）、载体、抗生素和大小（如果已知）信息	针对每个克隆集，提供克隆标识符、存取号，如果有的话，基因名称、基因符号、Unigene 类别 ID（当可适用时），包括 Unigene build 版本和染色体定位（如果已知）	需交纳年费以使用 arrayBASE，它是一个详尽的注释数据库，提供每个 arrayTAG 的深入信息。克隆发货时附有 QC 证书、FASTA 序列文件、注释摘要文件包含了多孔板坐标、GenBank 和 UG IDs，UG、bioSCOUT 描述
一般信息和计划等	—	资源中心（Resource Center）是经授权的 IMAGE. ESTs的 5 个发布者之一。IMAGE. ESTs 是目前可利用的最大的公开 EST 收藏，拥有超过三百万个克隆。另外，资源中心分发由德国 cDNA 联盟协定（S. Wiemann，DHGP）生产的克隆和哺乳动物基因收藏的克隆（*R. Strausberg*，NIH）	当前每个收藏中已经有 10 000 个克隆，很快会增加至 20 000 个非冗余的 cDNA 克隆集。cDNA 在一个相同的载体系统中重新克隆，以产生由 mRNA 的 3′端得来的相似大小的 cDNA。poly(A)尾部被去除，并且对每个序列标签进行重复的或低复杂性区域核查。利用内部的方法对序列进行聚类，并与 NCBI Unigene 聚类相匹配。利用 LION's bioSCOUT 序列注释系统，采用代表性的序列进行详细的克隆注释。所有的注释数据都能通过一个全面的，基于网络的注释数据库 arrayBASE 存取。关于更加详细的信息，请参见 www. lionbioscience. com/arrayTAG

a 网址：http://www. resgen. com

b 网址：http://www. rzpd. de

c 网址：http://www. lionbioscience. com

ATCC（www.atcc.org）也是IMAGE供应商之一，是在定购单个克隆时非常有价值的站点。为了找到感兴趣的克隆，最好使用一个以上的克隆ID来检索，因为有些克隆ID不再是唯一的；两个ID确保了选择的正确性，并且可以使定购过程加快。可以在GenBank中对克隆信息进行初步的检索来获取克隆ID。一旦确定了感兴趣的克隆，在ATCC网站的分子生物学部分使用一个ID检索，并用另一个来确认，应当很容易找到。如果没有能够在搜索一个特定的亚类中找到克隆，那么对整个分子生物学收藏进行全面的检索可能会得到结果。

42 信息栏：点样微阵列中使用的对照

Andrew Holloway（Peter MacCallum癌症研究所，东墨尔本，维多利亚，澳大利亚）

实验的变化可能发生在表达分析的任何阶段，包括从玻片点样到杂交，再到数据的获得。为了量化这些变化，许多研究者将一系列的对照DNA结合到阵列的单元中。分析从这些对照中获得的数据，为我们提供了关于杂交的动态范围、灵敏度和特异性的信息，并可以对观察到的表达比例的重复性进行统计学分析。通过在阵列的不同位置上点样许多种类的对照DNA，系统的空间变化和Cy3及Cy5信号强度的偏差可以被测量和校正。

对照的类型

如果条件允许，用制备其他阵列单元材料相同的方法来制备对照DNA，并以相同的方式点样。

阳性对照

这些对照用于和探针杂交，不论来源如何，被用来证实靶DNA被Cy3及Cy5标记上可接受的特异强度。每类cDNA（也就是待测和参比）与阳性对照杂交后应该产生大约相等的强度信号。将点样到阵列上各点的材料取等体积（约1 μl）合并，可产生一系列简单的阳性对照。做一系列稀释，如750 ng/μl、500 ng/μl、400 ng/μl、300 ng/μl、150 ng/μl、75 ng/μl以及10 ng/μl合并的DNA（pooled DNA）。全部的稀释系列应与合并DNA中的各种组分同时点样。杂交后，在很宽的浓度范围内，信号强度应该正比于每个对照点的DNA含量，并且待测样品和参比的表达比例也应接近于1（Yang et al. 2002）。

阴性对照

这些对照用来评价非特异性杂交的程度并由此设定一个信号阈值。阵列上任何一个信号不明显强于阴性对照的点都被认为没有结合上荧光标记的靶标。通常，阴性对照是来自与当前研究物种没有已知（或预期）的同源物或者共生同源物的生物中编码DNA的片段。用作哺乳动物基因表达分析的微阵列经常包含由合并的植物或细菌编码序列组成的阴性对照（Miki et al. 2001）。为了制备阴性对照，应选择一组与所研究的物种在

进化上相隔很远的基因或编码序列，但要包含相近的 G、C 含量。对照应以与阵列上其他样点相同的浓度进行点样。其他的阴性对照包括不含 DNA 的点样缓冲液和浓度为 4～40 pmol/L 的 poly（dA）$_{18\text{-}24}$溶液，用点样缓冲液配制。杂交和扫描完成后，核查从这些阴性对照中获得的信号在每个检测通道不显著高于背景水平。从阴性对照发出的信号通常表明探针与阵列的非特异性杂交，并可能证实需要在杂交和清洗时提高严格度。阴性对照可用来估计背景（见第 7 章）。

43

添加对照（spiked controls）

为了证实标记的 Cy3 和 Cy5 的量是相等的，向每个标记反应体系加入已知数量 poly（A)修饰的 mRNA 种类，这类 RNA 可根据编码链 3′端含有 poly（A）序列的重组克隆，使用噬菌体编码的 RNA 聚合酶按照标准方法在体外合成［如果 mRNA 是通过随机引物法标记，就不需要 poly（A）序列］。与用作体外转录模板相同的 DNA 片段也要制备和在阵列上点样。然而，点样材料中的 3′poly（A）序列应该略去。

添加对照应该与所研究的生物体或细胞系中任何已知表达序列无交叉杂交。通常，添加对照可从细菌基因、植物基因或者酵母基因间隔区（intergenic region）来合成。添加对照可以按照一系列浓度范围加入到待测和参比样品中，为系统的动态范围提供数据。在一系列浓度范围内，对应于这些对照的阵列单元上 Cy3 和 Cy5 探针杂交的强度应该相等。在含有 0.5 μg mRNA 或 50 μg 总 RNA 的标记反应中，添加对照的加入量（通常为 5 pg～25 ng）在数量上为内源性 RNA 量的 0.001%～5%。

一个互补的方法是系统地改变加入到两个标记反应体系中的添加靶标（spiking target）的量，然后把观察到的 Cy3∶Cy5 通道的信号比值与化学计算法的预测结果相比较。添加靶标要按 1∶10、1∶5、1∶2、1∶1、2∶1、5∶1 和 10∶1 的比率加入到两个标记反应体系。一般的，加入的添加靶标的浓度应位于系统动态范围的中间。

RNA 的完整性/反向转录对照

点样到阵列上的对照 DNA 也可以用来核查用作杂交的标记 mRNA 的完整性和反向转录的效率。在设计这些样点时，要从靶基因的 5′区域、中间区域及 3′区域产生一个基因的若干克隆。比较样品和这些点的杂交，得到在标记反应中复制到 cDNA 的靶序列长度信息。另外，如果来自靶基因 3′区域的样点与标记靶标，如 Cy3 和 Cy5 标记的靶分子具有相同的杂交效率，而来自 5′区域的样点总是显示 Cy3∶Cy5 的值大于 1，那么 Cy5 对远端序列的标记效率可能没有 Cy3 的标记效率高。可以通过交换标记来证实。

标准化

这种对照通常用于数据的标准化，也就是说，消除微阵列实验中影响基因表达水平测量结果的系统偏差。例如，由于两种荧光染料标记效率的不同将会导致系统偏差的产生。Yang 等（2002）和网站 http://www.stat.Berkeley.edu/users/terry/zarray/html/normspic.html 中描述了可用于标准化的三种基因类型。

44

• 包含在阵列上基因中的 DNA

在许多类型的微阵列实验中，被比较样本中只有一小部分基因预期会有差异性表达，其余那些被认为表达水平没有改变的基因，可用于对来自两种染料的信号强度的标准化。使用内源性 DNA 作为标准化对照的优点是阵列中基因表达水平之间的非系统性差异可以通过对大量阵列样点的红色和绿色信号的平均来平整。

- 恒定表达的基因

  一些研究人员在微阵列上点样少量的被认为在众多细胞中表达水平恒定的持家基因。这类基因如亲环素（cyclophilin）、甘油醛-6-磷酸脱氢酶（G6PD）及β-肌动蛋白。然而，无法保证被用作标准的持家基因在不同样本之间的表达保持不变。事实上，有数据显示常用作标准化的基因的表达水平在不同的哺乳动物组织或细胞系中或者在不同生长条件下的同种细胞中并不是恒定的（Mansur et al. 1993；Spanakis 1993；Bhatia et al. 1994；Eickhoff et al. 1999）。

- 添加对照

  如先前所述，添加 DNA 对照用于点样到微阵列上（通常重复点样）。通常添加 DNA 由与当前研究物种无关的生物体的 DNA 序列组成。标记之前，对应的 RNA 在两份 RNA 制备物中的含量相同。在没有系统差异时，红色和绿色通道的输出结果应该是相等的。有些研究人员点样一系列添加对照含量不同的阵列样点，预计在这些浓度范围内红色和绿色荧光的强度应该是一样的。

Yang 等（2002）和 http://www. stat. Berkeley. edu/users/terry/zarray/html/normspie. html 以及第 7 章中描述了在表达分析实验中，对来自外源和添加对照的数据进行标准化的统计学方法。

商品化对照集

一些公司销售含有用于点样的 DNA 样品的试剂盒，同时提供配套的 poly（A）$^+$ RNA（添加混合物），在鼠和人的基因表达分析中用作验证、标准化和对微阵列的性能的监测。

- Amersham Pharmacia Biotech 销售的 Lucidea Microarray ScoreCard（如人、RPK1161B）用人工合成的基因和 mRNA 来测量动态范围，提供了一个软件系统用于数据的分析和作图。关于更详细的信息，请参见 http://www. apbiotech. com/lsn _ home/lsn _ index. html。
- Stratagene Inc.（252005）销售的 SpotReport，用三种拟南芥（*Aravidopsis thaliana*）基因的 PCR 产物和相应的 poly（A）$^+$ RNA 来测量动态范围和建立比率对照。关于更详细的信息，请参见 Davis 等（2001）（可以从 http://www. stratagene. com/vol13 _ 4/p128-130. htm 在线获取）。

这两种商品化的对照集中都包含了持家基因和多种可用作阴性对照的 DNA。

45 信息栏：基因组和 cDNA 文库中 T1 噬菌体的预防和检测

Norma J. Nowak（布法罗纽约州立大学医学与生物医学科学学院，纽约布法罗 · 罗斯威尔公园癌研究

所，微阵列和基因组研究室主任）

噬菌体 T1 污染易受感染的细菌保藏物的事例很多。正如 Drexler（1988）所指出的“……它短暂的潜伏期和抗干燥能力给细心的工作人员带来了挑战，而对无菌技术不佳的人而言则是一个噩梦”。作为人类基因组计划（HGP）的一部分而建立起来的有价值的基因组和 cDNA 文库，绝大部分是在抗 T1 噬菌体的大肠杆菌菌株可以获取之前构建的。

许多由 HGP 支持的大规模作图和测序项目需要平行培养数千个易受 T1 噬菌体感染的培养物。在这一过程中，尽管绝大多数微生物污染物可以得到控制，但 T1 却能够很快传播，并且有时使易受感染的培养物几乎无法克隆或繁殖。通常，最初的污染是不被察觉的。一小部分克隆培养物在微滴定板上生长失败被认为是接种不成功所致。然而，有时候导致这一现象的原因是难以捉摸的，表面上没有生长是由于 T1 溶菌污染的结果。噬菌体一旦在实验室中生长，便能够通过气溶胶迅速传播，感染相邻的微滴定板孔并毁坏有价值的文库。可以采取一些测量方法来保护易受感染的克隆组，并监控 T1 的污染。

顶层琼脂分析（top agar assay）检测噬菌体

采用高通量分析技术的实验室涉及大量的克隆保藏品，需要用下述的顶层琼脂分析方法对克隆集进行常规检测。该分析方法能够用于检测单个培养物，也可以用于快速证实微量滴定板上可疑的 T1 污染。

材料

缓冲液和溶液

储存液、缓冲液和试剂的成分见附录 2。

将储存液稀释至合适浓度。

水（灭菌）

次氯酸钠（10％和 20％，漂白剂）或 10％的 Tegedyne

培养基

LB 培养基

LB 琼脂平板（薄层，22 cm×22 cm；Nunc），不含抗生素

在琼脂薄板上的 LB 琼脂层约 2.5 mm 深，大致为标准厚度的一半。在每块 22 cm×22 cm 平板上倾倒约 120 ml LB 琼脂。平板可以在 4℃下存放，倒置在塑料盒中，最多 2 天。

LB 顶层琼脂（500 ml 的 LB＋0.8％琼脂，装在瓶中）

菌株

DH10B（Invitrogen）

专用设备

细菌平板（22 cm×22 cm；Nunc）

振荡培养箱，设定为 37℃

水浴，设定为 42℃

方法

1. 进行噬菌体检验的前夜，接种 DH10B 细胞于 50 ml 的 LB（不含抗生素）中。将培养物置于振荡培养箱中，37℃下培养过夜。
2. 用保温的新鲜 LB（150～200 ml）稀释过夜培养物至 OD_{600}＝0.25。在 37℃振荡培养箱中，孵育培养物 1 h，直到 OD_{600}＝0.45。这一阶段，培养物处于生长的对数中期。
3. 从培养箱移去 DH10B 的培养物，在室温下保存备用。
4. 将 50 ml 瓶中的 LB 顶层琼脂在微波炉中熔化，然后置于 42℃的水浴中至少 2 h。每 30 min 转动溶液一次，确保熔化的琼脂胶均匀混合和冷却。

   注意：42℃是一个关键的温度。如果琼脂胶太热，将会杀死 DH10B；如果太冷，溶液会过早胶化。
5. 取出存放在 4℃的 LB 琼脂平板，保存在 37℃下直到准备倾倒顶层琼脂（最少保存时间为 30 min）。
6. 在前后彻底清洁的层流通风橱中，用 10％漂白剂或者 10％ Tegedyne 灭菌，然后用无菌水洗净。
7. 将琼脂薄板排放于工作台上，在平板底部和上盖做标记，然后移走上盖。
8. 从水浴中移走装有熔化的顶层琼脂的 500 ml 瓶，并加入 11.25 ml 处于对数中期的 DH10B 培养物。轻轻摇摆瓶体将细胞悬液完全混合，防止产生气泡。
9. 向 LB 琼脂薄板上倾倒一层已接种的顶层琼脂（约 80 ml）。用灭菌的移液器吸头将气泡赶至平板边缘。

   注意：不能对平板火焰灭菌，因为这会杀死 DH10B 细胞！
10. 让平板冷却。在将上盖放回板上之前要确保顶层琼脂已经完全固化。

    注意：平板必须在制备当日使用或丢弃。

    一旦噬菌体分析平板冷却和固化后，就可以使用 96 针或 384 针复制器转印生长在微量滴定板中的培养物样品，或者划线培养单个克隆。
11. 转印后的平板，**不要翻转**，在 37℃下孵育过夜，用 Saran Wrap 和 Parafilm 密封。

    接种在顶层琼脂中的 DH10B 细胞生长过夜，使得琼脂层变得不透明。如果在待测样品中存在噬菌体，它将感染 DH10B 细胞，导致在 DH10B 菌苔上出现清晰的噬菌体噬斑。菌苔上"清晰斑点"非常明显，容易识别（图 1-4）。被细菌感染后消除了包围在检测时划线克隆周围的 DH10B 菌苔。
12. 将剩余的 DH10B 培养物和装有顶层琼脂的瓶灭菌。

    剩余的 DH10B 培养物：加入等体积 20％的漂白剂并彻底混合溶液，放置至少 15 min，然后

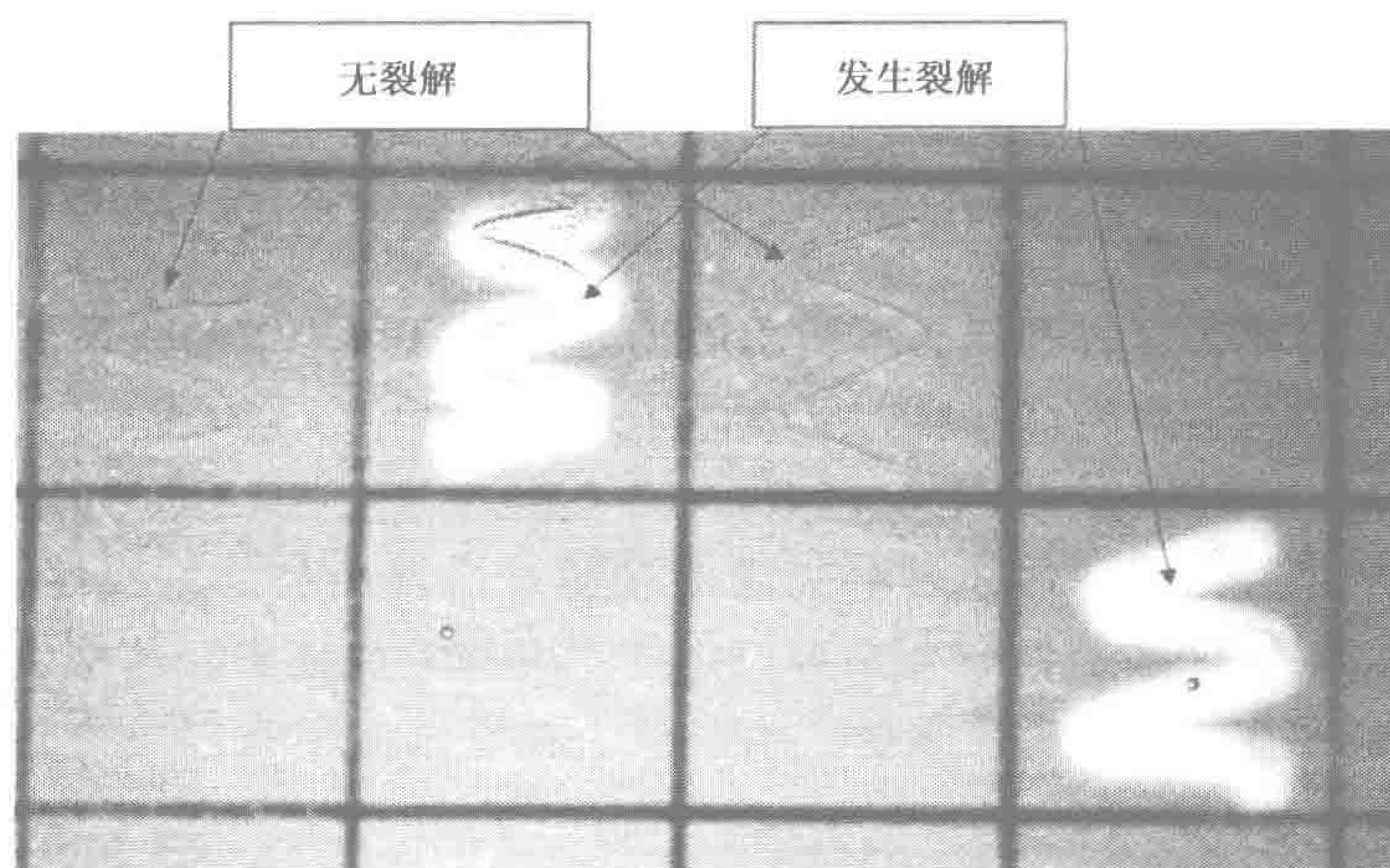


图 1-4　T1 污染的检测

噬菌体分析平板示例了噬菌体感染的 BAC 克隆培养物经过 18 h 孵育后，顶层琼脂上菌苔的裂解

47

用大量自来水冲洗溶液到水槽中。

装有顶层琼脂的瓶：向瓶中灌入 10%漂白剂并剧烈摇动瓶中溶液，放置至少 15 min，然后用大量自来水冲洗溶液到水槽中。

预防 T1 噬菌体污染

时刻保持完善的无菌技术可以防止 T1 噬菌体的污染。采用如下措施可以控制克隆集中污染的扩散。

- **在抗噬菌体的宿主中构建基因组和 cDNA 文库**。预防 T1 噬菌体感染的最好方法是在抗噬菌体的宿主，如 MAX Efficiency DH5α 细胞、Electromax DH10B 抗-T1 噬菌体细胞或"Gene Hogs"（Invitrogen）中构建基因组和 cDNA 文库。有抵抗力的细胞携带了一个 FhuA 膜蛋白（tonA）突变（Killmann et al. 1996）。通常用于铁吸收的 FhuA 蛋白，含有一个门控环（gating loop），用作 T1 及其近亲 T5 感染粒子的受体（Killmann et al. 1995）。携带了一个 FhuA 突变的 DH10B 细胞使得它们能够部分抗噬菌体 T5，并且完全抗 T1 感染。通过酚：氯仿抽提从每个克隆分别制备质粒 DNA，随后重新转化到抗噬菌体的宿主中，可以挽救或者恢复被污染的克隆组。

48

- **减少受感染平板中 T1 的滴度**。在含有 1 μmol/L 铁色素（F-8014，Sigma）的培养基中培养细胞，受感染的微量滴定板上活的噬菌体 T1 的滴度能够有效地减少。铁色素结合到 FhuA 受体上，阻断了噬菌体 T1、T5 和 φ-80 的结合。虽然这种方法不能杀灭噬菌体，但可以显著地减少滴度，直至有价值的克隆保藏品能被挽救。这种方法是作为伯克利果蝇基因组计划（Berkeley *Drosophila* Genome Project）的一部分发展而来（见 http://www.fruitfly.org/dgc/phage.html）。
- **在层流通风橱中操作克隆和文库**。必须始终在层流通风橱中操作克隆和文库。这包括挑选、划线培养单克隆、转移培养物、排布克隆阵列、噬菌体检测和其

他操作。使用通风橱时，要避免堵塞有筛网的空气循环系统。

- **工作区域灭菌**。层流通风橱使用前后，必须对工作区域灭菌。用10%的漂白剂或者10% Tegedyne处理，接着用无菌水清洗。使用过的移液管和镊子也要用10%的漂白剂灭菌。
- **保持通风橱中灭菌灯开启**。不使用层流通风橱时，其中的灭菌灯必须开启。整个晚上和周末均要将灯开启，早晨将灯关掉。

  注意：不要观看灭菌灯或其反射光线，否则在数秒内会对眼睛造成严重损伤。
- **使用废物烧杯**。任何用于转移克隆或接触过活培养物的一次性吸头或者牙签，必须放入衬有填塞了10%漂白剂口袋的“废物”烧杯中。这些吸头必须放在通风橱中，并用灭菌灯整夜照射。第二天，在通风橱中将口袋密封，置于消毒包中焚化。
- **消毒培养箱**。使用后，必须用10%的漂白剂擦洗振荡培养箱，接着用无菌水清洗。
- **使用天花板上的灭菌灯**。实验室应该在天花板上安装灭菌灯，通过一自动定时器在半夜实验室无人时开启3h。灯开启之前要发出警铃。当灯在使用中时，实验室入口处警灯要打开。

  注意：不要观看灭菌灯或其反射光线，否则在数秒内会对眼睛造成严重损伤。
- **使用便携式灭菌灯**。便携式灭菌灯用于操作和处理细菌培养物区域的灭菌，包括振荡培养箱和有害生物废物的容器。

49　信息栏：长链寡核苷酸的点样：替代用于表达分析的cDNA阵列

使用cDNA克隆集制备阵列的缺点是大量克隆的具体操作，包括复制和扩增，需要消耗大量的劳力，同时所包含处理的复杂性还可能导致错误的出现。相对于cDNA衍生的PCR产物阵列，用合成的长链寡核苷酸（>50 bp）集点样制备的微阵列代表了有吸引力的替代方法。使用商业化的寡核苷酸集避免了微阵列制备前材料准备过程的许多工作，同时也利用了快速增长的基因组序列信息。另外，可以选择代表特定可变剪接体的序列，避免相关基因的交叉同源区域，5′端氨基修饰能使寡核苷酸与基底定向结合。理论上，研究人员可以自由合成任何序列的寡核苷酸集，而且不受具体克隆能否获取的限制。然而，由于成本和获取的容易程度等原因，许多研究者选择商业化的寡核苷酸集，它们包括大量由最新构建的相关基因组序列确定的基因（表1-6）。然而，对于某个具体项目，要对商业集增添特异性寡核苷酸。初期研究结果表明，从PCR扩增的cDNA点样的阵列与长链寡核苷酸微阵列两者得到的结果具有高度的一致性（Hughes et al. 2001）。事实上，商业文献中报道从酵母基因组寡核苷酸集（yeast genome oligo set）（Operon）获得的结果，比平行实验中扩增cDNA获得的结果更为准确，重复性更好。随着基因组DNA序列分析的复杂性和精确性的增加，对应于单个外显子的长链寡核苷酸或许最终可以取代目前占主导地位的扩增的cDNA。然而，这些目前还不成熟。关于新开发寡核苷酸集的信息较少，研究人员从cDNA过渡到长链寡核苷酸微阵列时必须仔细考虑。

仔细设计的寡核苷酸是其良好性能的关键。使用喷墨打印法合成一系列人和酵母寡核苷酸集，Hughes 等（2001）评价了影响灵敏度和特异性的寡核苷酸设计的关键。长度为 60 bp 的寡核苷酸，在 32%甲酰胺存在下，40℃时与靶序列复性，杂交获得了最佳的灵敏度和特异性。为了获得最高特异性，一定要避免高含量的脱氧胞嘧啶以及相关基因在序列上的相似性。

信息栏末的表 1-6 中列举了一些用于表达分析的商业化寡核苷酸集，包括关于它们的设计和质量控制的背景信息。设计上共同的特点有：

- **选择正确的（反义）链并且优先在 mRNA 的 3′区域序列作图**。选择正确的链是最基本的，然而，正如在 Affymetrix 小鼠寡核苷酸阵列中发生的令人尴尬的情形那样（Knight 2001），这一点常常会被忽略。由 mRNA 的 3′-UTR 衍生的序列单元不太可能与基因家族中成员产生交叉杂交。另外，大多数标记靶序列的合成中使用了 oligo（dT）作为引物，应用对应于 3′序列的探针可使由截短的靶序列引起的问题减至最小。截短可能是由于 mRNA 模板的降解，尤其是从组织获得的模板，或者是在掺入大的荧光标记核苷时因空间位阻造成的低效率反转录。
- **选择通过分析大量 EST 确定的一致性序列以避免序列差错**。序列差错引起的单碱基错配对长度大于 50 bp 的寡核苷酸的杂交性能影响不大（Hughes et al. 2001）。然而，在靶和探针序列之间出现几个碱基错配将会降低杂交特异性和杂交动力学。影响程度的大小取决于两条序列之间错配的位置。
- **选择位于窄的熔解温度（T_m）范围内的寡核苷酸**。由于寡核苷酸集中所有的寡核苷酸探针具有相同的杂交和洗涤条件，因此所有寡核苷酸具有近似的熔解温度就显得非常重要。可以通过选择具有相近 G+C 含量的序列和（或）应用一种可使用的算法计算全部寡核苷酸理论上的 T_m 值来实现（Sambrook and Russell 2001）。
- **避免容易交叉杂交或者自我互补的序列**。交叉杂交限制了分析的特异性。重复性或者低复杂性的序列可以通过在寡核苷酸集中筛选已知序列单元的方法加以排除。这些序列中应该没有大于 3 bp 的反转或者自我互补序列出现，因为这些序列能形成发夹结构，减少靶序列杂交的特异性。最后，应用如 BLAST 等用于大规模序列比较搜索的工具，可以在寡核苷酸集中筛选与衍生靶分子的基因组序列之间的偶然同源性。
- **选择代表具有有限或者至少已知冗余度的单个基因的寡核苷酸**。转录基因的识别可以采用外显子预测程序，如 ORF 搜索软件（http://www.ncbi.nlm.nih.gov/gorf/gorf.html）和 EST 信息。针对内含子和重复序列非常丰富的复杂哺乳类动物或者植物基因组的大多数寡核苷酸集是使用 EST 数据设计的。正如关于**克隆集**的信息栏中所描述的，对集中于 Unigene 和相似数据库中 EST 数据的不断修订，导致一些类别（cluster）随着时间的推移发生坍塌。因此，用于提供寡核苷酸合成信息的特定 Unigene“构建版本”非常重要，同时能为寡核苷酸集可能代表的独立基因的数量给出一些提示。Hughes 等（2001）主张一条经过仔细挑选的寡核苷酸足以代表每一个基因，尽管一些研究人员认为每个基因使

50

用两条或三条寡核苷酸为实验结果提供了重要的内在验证。Affymetrix 公司的寡核苷酸微阵列是这种方法的一个极端例子，每个“待测基因”使用了约 20 条寡核苷酸。然而，这种情况下寡核苷酸链实际上是很短的（20～24 目），因而与大于 50 bp 的寡核苷酸相比，杂交特异性较差，杂交条件也更为讲究。

商业化寡核苷酸设计的确切方法通常是商业秘密。然而，作为一个例子，Hughes 等（2001）描述了一个逐步进行的寡核苷酸设计方法。

1. 分别选择位于酵母和人基因转录物 3′端 1 kb 或 50～350 bp 范围内的 60mer 序列。
2. 低复杂性和重复性序列应用 RepeatMasker［华盛顿大学基因组中心，（http//www. genome. Washington. edu/uwgc/analysistools/repeatmask. htm）和遗传信息研究所（http//www. girinst. org)］软件进行掩蔽。RepeatMasker 软件不会评价自身互补性，除非在序列中出现回文重复（palindromeic repeat）。
3. 选择 35%～45% G+C 含量的探针（酵母），或者小于 50% G 含量并且 G+C 含量为 20%～70% 的探针（人），放弃含有超过 6 个碱基的均聚延伸片段的序列。这一点对于获得具有窄的熔解温度范围的寡核苷酸来说极为重要。
4. 与已知基因具有同源性的 EST 链的定向通过使用数据库中的注释获得。采取的其他方法是基于有义和反义引物杂交动力学的比较，其中正确定向的寡核苷酸的特异性杂交具有比较缓慢的杂交动力学特征。
5. 通过这些筛选步骤的寡核苷酸用 BLAST（http//www. ncbi. nih. gov/BLAST）与各自的酵母和人数据库比对，以验证序列的相似性，选择最具有特异性的序列用于合成。

商业化的软件包，如 Array Designer（http//www. premierbiosoft. com/dnamicroarray/dnamicroarry. html）可用于微阵列实验中的自动化设计。

信息栏：酵母探针

Saccharomyces cerevisiae（面包酵母、啤酒酵母或者芽殖酵母）是第一个全基因组序列被全部测定的真核细胞（Goffeau et al. 1996），同时，它也是第一个全基因组表示在 DNA 微阵列上的生物体（DeRisi et al. 1997）。这些和随后的酵母基因表达研究中，采用由代表整个完备的酵母基因的全长 PCR 扩增的序列，通过点样制备 DNA 微阵列。

在许多方面，以这个模式生物体来设计和制备基因表达的 DNA 微阵列代表了一种“容易”的情形。已知的和预测的蛋白质编码序列产生共计约 6000 个基因；编码序列的注释是相当充足的；而且通过增加少量表达分析技术，如 SAGE 等确认的基因后，基因目录就完成了（Velculescu et al. 1997）。酵母中的可读框（open reading frames，ORF）相当的小，平均大小为 1.1 kb，并且 80%的 ORF 在长度上是 2 kb 或者更短。另外，非常少的（6000 个基因中约 230 个，或者 3.8%）转录子经过剪接。基因组的这些特性对以酵母设计和制备微阵列有很大的影响。基因目录完成不久以后，以酵母基因组 DNA 作为模板，用于扩增在已知的或预测的启动和终止密码子之间全部序列的引物对，很快便可以购买得到（Research Genetics，http://www. resgen. com)。这些引物

通常包含限制性酶切位点，以便将 PCR 产物克隆到标准的质粒载体中，同时也具有共同的 5′端，允许只需要一对两侧引物（a single pair of flanking primers）复性到共同的 5′端就可以重新扩增。

除了表达分析以外，使用基因错配扫描技术的 DNA 微阵列已经应用于酵母中遗传上相同（identicall-by-descent，IBD）位点的作图（Cheung and Nelson 1998），直接比较 *S. cerevisiae* 的相关变种的基因组，以及在全基因组范围对启动子上 DNA 结合蛋白的在体结合的分布进行作图（Lyer et al. 2001；Reid et al. 2000；Ren et al. 2000）。这是因为紧靠一个酵母基因的上游基因间隔区域中包含了控制其表达的元件，含有这些元件的基因间隔区的 DNA 片段可以通过 PCR 扩增，并用于制备 DNA 微阵列。

52

将酵母序列表示到 DNA 微阵列上的另一种替代方法是使用长链的寡核苷酸（60～70 mer），它们具有容易制备的优点，并且避免了保守的编码区域之间的交叉杂交。然而，PCR 扩增片段可能更适合于分析基因间隔区域，例如，先前就没有关于蛋白质结合位点的知识。

已经测序的 *S. cerevisiae* 基因组是来自 S288c 菌株（*MATα mal gal2*）。来自该菌株的基因组 DNA 通常用作扩增的模板，配套的基因组的 PCR 引物对只适合于这一菌株。如果由 S288c DNA 序列制备的微阵列用于相关酵母菌株的基因组比较在其他菌株中出现，而在 S288c 中缺失的基因组位点将不能被证实。

Research Genetics 公司合成了第一套用于扩增所有酵母基因的 PCR 引物对集，在本书写作时，该公司仍然是该引物集的唯一商业来源。Research Genetics 有多种其他引物对可供选择（表 1-3）。用于扩增整个 ORF 的引物对称为 GenePairs，整个引物集有两种不同的浓度可供选择。有可以用于扩增基因间隔区域的引物对，也可以获得两种不同的浓度。还有一种引物对（pORF Primers）集，用于扩增一部分的 ORF。与标准 ORF 引物集相比，这种引物对的优点是产生的 PCR 产物更短一些，所有产物的长度均在 1 kb 以内。这主要是为了减少一些与通过 PCR 进行较长的 ORF 全长扩增相关的问题。另外，pORF 是与 ORF 的非保守区域结合，用以减少在序列高度相似基因中可见的交叉杂交水平。pORF 集的缺点是引物没有一致的侧翼序列，因而无法像标准 ORF 引物集中那样，使用一对共同的引物对 PCR 产物再扩增。

可获得的两种不同浓度的基因间隔区引物对集是基因组作图和染色质免疫沉淀（chromatin immuneoprecitation）所必需的（第 6 章方案 10 和方案 11）。基因间隔区引物集与 ORF 集一起，几乎覆盖了酵母基因组的每一个位点。目前，来自核基因组的非 ORF 特征序列（non-ORF features），如 rRNA、tRNA、snoRNA、Ty-LTR 元件等和所有的线粒体基因，都包含在“基因间隔区的”引物集当中。

虽然酵母基因组序列特征的注释已经取得了相当大的进展，但它仍然是一项正在进行中的工作。许多 ORF 的性质完全没有了解，并且它们当中的一些有可能是不被转录的。虽然这些片段也将包含在 ORF 集内，正确地说它们属于基因间隔区集。同样地，一个看似没有特征的基因间隔区片段或许携带了一种秘密的转录物，可能是编码一种小分子蛋白质或者 RNA。因此，研究人员必须有将一些新成分引入到微阵列中的准备；如果使用完整的集，要能够对档案数据库中现有成分的注释进行更改。

表 1-6 寡核苷酸集供应厂商

	Compugen[a]	Operon[b]	MWG Biotech[c]
寡核苷酸文库集	**人类寡核苷酸文库** 寡核苷酸数目:18 861 代表的 LEADS 人的聚类类别:18 664(加 197 个对照) 代表的 UniGene 人的聚类类别:15 223 在 UniGene 数据库没有能够很好辨识的公共域人类基因: 3441 每个 LEADS 聚类类别的寡核苷酸数目:1 寡核苷酸长度:60 bp(5′-C6 氨基修饰) 与 3′端的距离:95%的寡核苷酸与 3′端距离为 61~1309 bp。与 3′端的平均距离为 397 bp 交叉同源性: 95%的寡核苷酸预测的交叉同源性小于66.2%。平均交叉同源性为 33.8% 人类 UniGene 版本 #138 用于将寡核苷酸和 UniGene 聚类类别对应	**人类基因组寡核苷酸集** 寡核苷酸总数:21 329 代表的人类基因数目:21 329(加 30 个对照) 寡核苷酸长度:70 bp(5′-C6 氨基修饰) 与 3′端的距离:距离 3′端在 750 bp 以内。 Human Genome Oligo Set 的所有寡核苷酸根据 UniGene 数据库,版本 Hs #147 中具有代表性的序列设计	**泛人类 10K 寡核苷酸集** 寡核苷酸数目:9984 代表的 UniGene 聚类类别:9850 每个 UniGene 聚类类别的寡核苷酸数目:1 寡核苷酸长度:50 bp(5′-C6 氨基修饰) **泛人类癌症寡核苷酸集** 寡核苷酸数目:1920 代表的 UniGene 聚类类别:1900 每个 UniGene 聚类类别的寡核苷酸数目:1 寡核苷酸长度:50 bp(5′-C6 氨基修饰)
	小鼠寡核苷酸文库 寡核苷酸数目:7524 代表的 LEADS 小鼠的聚类类别:7445(加 79 个对照) 代表的 UniGene 小鼠的聚类类别:6544 在 UniGene 数据库没有能够很好辨识的公共域小鼠基因: 901 每个 LEADS 聚类类别的寡核苷酸数目:1 寡核苷酸长度:65 bp(5′-C6 氨基修饰) 与 3′端的距离:95%的寡核苷酸与 3′端距离为 65~1133 bp。与 3′端的平均距离为 396 bp。 交叉同源性: 95%的寡核苷酸预测的交叉同源性小于66.5%。平均交叉同源性为 30.5% 小鼠 UniGene 版本 #93 用于将寡核苷酸和 UniGene 聚类类别对应	**小鼠基因组寡核苷酸集** 寡核苷酸总数:16 463 代表的小鼠基因数目:16 463(加 30 个对照) 寡核苷酸长度:70 bp(5′-C6 氨基修饰) 与 3′端的距离:距离 3′端在 750 bp 以内 Mouse Genome Oligo Set 的所有寡核苷酸根据小鼠 UniGene 数据库,版本 Mm #102 中具有代表性的序列设计	**泛小鼠 10K 寡核苷酸集** 寡核苷酸数目:9984 代表的 UniGene 聚类类别:9850 每个 UniGene 聚类类别的寡核苷酸数目:1 寡核苷酸长度:50 bp(5′-C6 氨基修饰) **泛大鼠肝脏寡核苷酸集** 寡核苷酸数目:1408 代表的 UniGene 聚类类别:1353 每个 UniGene 聚类类别的寡核苷酸数目:1 寡核苷酸长度:50 bp(5′-C6 氨基修饰) **泛大鼠 5K 寡核苷酸集** 寡核苷酸数目:5760 代表的 UniGene 聚类类别:5535 每个 UniGene 聚类类别的寡核苷酸数目:1 寡核苷酸长度:50 bp(5′-C6 氨基修饰)

	Compugen[a]	Operon[b]	MWG Biotech[c]
寡核苷酸文库集	**大鼠寡核苷酸文库** 寡核苷酸数目:4854 代表的 LEADS 大鼠的聚类类别:4803(加 51 个对照) 代表的 UniGene 大鼠的聚类类别:3821 在 UniGene 数据库没有能够很好辨识的公共域大鼠基因:982 每个 LEADS 聚类类别的寡核苷酸数目:1 寡核苷酸的长度:65 bp(5′-C6 氨基修饰) 与 3′端的距离:95%的寡核苷酸与 3′端距离为 65～928bp。与 3′端的平均距离为 302 bp 交叉同源性:95%的寡核苷酸预测的交叉同源性小于 66.2%。平均交叉同源性为 29.2% 大鼠 UniGene 版本 #90 用于将寡核苷酸和 UniGene 聚类类别对应	**酵母基因组寡核苷酸集** 寡核苷酸总数:6307 代表的 *S. cerevisiae* 的 ORF 数目:6307(加 10 个对照) 寡核苷酸长度:70 bp(5′-C6 氨基修饰)与 3′端的距离:距离 3′端在 750 bp 以内 该寡核苷酸集代表了从斯坦福大学 *Saccharomyces* 基因组数据库中得到的 *S. cerevisiae* 的 ORF	**泛酵母寡核苷酸集** 寡核苷酸数目:6368 代表的 UniGene 聚类类别:6250 每个 UniGene 聚类类别的寡核苷酸数目:1 寡核苷酸长度:40 bp(5′-C6 氨基修饰)
		疟疾基因组寡核苷酸集 寡核苷酸总数:6369 代表的 *Plasmodium falciparu* 的基因数目:6231(加 138 个对照) 寡核苷酸长度:70 bp 与 3′端的距离:距离 3′端在 750 bp 以内 该寡核苷酸集代表了来自 *P. falciparu*的 ORF *P. falciparu* 的基因组序列来自桑格中心(Sanger Center)、TIGR 和斯坦福大学。根据该基因组序列,应用 TIGR 开发的基因预测软件,用于疟疾的 GLIMMR,预测 ORF	**泛 *H. pylori* 寡核苷酸集** 寡核苷酸数目:1920 代表的 UniGene 聚类类别:1877 每个 UniGene 聚类类别的寡核苷酸数目:1 寡核苷酸长度:40 bp(5′-C6 氨基修饰)
		结核基因组寡核苷酸集 寡核苷酸总数:4321 代表的 *Mycobacterium tuberculosis* 的基因数目:4269(加 26 个对照) 寡核苷酸长度:70 bp 与 3′端的距离:距离 3′端在 750 bp 以内大部分寡核苷酸根据桑格中心从 *M. tuberculosis* 基因组预测的3924个编码蛋白质的基因序列集合设计。这个预测的基因序列集合来自 H37RV 菌株(实验室菌种),已在桑格中心全部测序。一小部分寡核苷酸集根据由 TIGR 测序和预测的菌株 CDC-1551 设计。要注意已经有 371 个来自该菌株的预测基因被发现与 H37RV 的预测基因有小于 97%的同源性,并且这些基因也包括在寡核苷酸集当中	**泛 *E. Coli* K12 寡核苷酸集** 寡核苷酸数目:4416 代表的 UniGene 聚类类别:4289 每个 UniGene 聚类类别的寡核苷酸数目:1 寡核苷酸长度:40 bp(5′-C6 氨基修饰)

54
55

续表

	Compugen[a]	Operon[b]	MWG Biotech[c]
寡核苷酸的设计	通过对 GenBank EST 与 mRNA 序列的聚类和拼接，产生这些寡核苷酸文库（OligoLibraries）（设计中使用了 GenBank 版本 #121）。聚类算法考虑了已知的重复单元、低复杂性区域、嵌合序列（chimeric sequence）、载体污染和正确的序列定向。寡核苷酸序列经过优化，使得与 3′端的距离、交叉杂交和二级结构最少，并且整个集合中 GC 含量的均一性以及 T_m都标准化 拼接算法对序列错误、预测的反义基因进行了校正，并为每个基因根据表达的和基因组数据去设计其他可选择的剪接方式和预测转录子。使用这些标准，每条寡核苷酸选自一条共有的序列片段，在该基因预测的转录物中包含这个序列片段的转录物数量最多	由 Operon 开发的特有的程序用于设计长度为 70 bp 和具有标准化熔解温度，即 T_m 在±2℃范围的寡核苷酸。计算 T_m的公式如下： T_m=81.5+16.6×log[Na^+]+41×(#G+#C)/序列长度−500/序列长度，其中[Na^+]=0.1mol/L。 考虑到由 oligo(dT)作为引物的不完全反转录反应，寡核苷酸设计为与距离基因（ORF）的 3′端 750 bp 范围内杂交。通过利用 BLAST 对包括 UNIGENE 基因组数据库在内的大量数据的搜索，序列经过优化与其他基因具有最小的同源性（因此交叉杂交最小）。这种设计可以区分重叠基因或选择性剪接的基因	根据 3′端序列和占优势的编码序列来选择寡核苷酸。属于同一基因的冗余序列条目进行拼接，最完全的编码序列录入 MWG CodeSeq 数据库中。CodeSeq 数据库包括相同基因的选择性剪接形式。从每个物种一特定种类基因索引可以搜索到感兴趣的基因。同样的搜索可以知道一个给定基因归属的家族，以及用哪些寡核苷酸去检测它，该基因的选择性剪接形式或者家族成员。根据基于 BLAST 和 Smith-Waterman 搜索的广泛比对，检查了寡核苷酸的交叉杂交。通过检验反向互补使二级结构最小化。对照 EST 和基因组数据库（使用的 UniGene 数据库是最新发布的），使用 BLAST 优化了序列。MWG CodeSeq 数据库用于全部的核查 在计算机设计中，寡核苷酸根据物理参数进行选择，包括 GC 含量、T_m、二级结构、引物二聚体形成和独特的结合位点
寡核苷酸的合成	OligoLibraries 由 SigmaGenosys 利用特有的 Abacus 寡核苷酸合成平台制造，耦联效率为 98.5%。对人序列而言，全长序列最低含量为 41%，而小鼠和大鼠的为 38%	小鼠、人和酵母基因组的寡核苷酸集均带有 5′-C6 氨基修饰 耦联效率为 99%时，可以得到约 50%的全长序列产物，而耦联效率为 99.5%则可以得到约 70%的全长序列产物。达到的效率通常较高，为 99.25%或者更好。合成中观察到的封闭效率接近于 100%，其结果是几乎没有中间碱基缺失的产物（报道的突变率为 1/1000）	所有寡核苷酸应用特有的 HPSF 技术合成，确保了高百分比的全长寡核苷酸，并且无污染（盐、金属等）。寡核苷酸均为带有 5′-C6 氨基修饰的 50 mer（适合的时候，一些非哺乳动物的阵列使用 40 mer）
质量控制程序	质量控制包括每条寡核苷酸的光密度（OD）测量。每批合成的寡核苷酸通过 MALDI-TOF 质谱仪检测以证实其分子质量和确保寡核苷酸在多孔板上的正确位置及序列的正确性	通过对三苯甲基峰、最终的三苯甲基峰和碘峰检测来评价合成过程是否正常运行。碘峰和三苯甲基峰是对所有寡核苷酸一起测量，而不是分别测量单个寡核苷酸。为每条寡核苷酸检测最终的三苯甲基峰，用来提供最后一步耦联反应的信息。用紫外（UV）光谱对寡核苷酸作常规评价。（70 mer 的寡核苷酸质谱无法检测）。用一部分寡核苷酸进行毛细管电泳来评价全长寡核苷酸的产量。紫外光谱仪用于对偶联效率的直接评价，因为如果偶联效率下降 1%或者更多，全长序列的产率将急剧下降	所有寡核苷酸集均进行三苯甲基监测和 MALDI-TOF 质谱仪分析以确保 90%的纯度。条形码和产品跟踪保证了制成多孔板的产品 100%得到身份确认

续表

	Compugen[a]	Operon[b]	MWG Biotech[c]
供货形式	寡核苷酸文库集以96孔板或者定制的386孔板形式提供。寡核苷酸文库的子集(subset)以单个的384孔板或者若干组每组包含了4块96孔板的形式提供。为了便于选择来自功能相似基因的寡核苷酸，寡核苷酸在多孔板上的排列依照标准的Gene Ontology (GO，见http://www. geneontology. org/)的指派 浓度：50 μmol/L(水中)或者1 nmol 推荐的点样浓度：20 μmol/L 推荐的点样表面：全部主要的表面	基因组寡核苷酸集以每孔2000 pmol的96孔板，或者每孔600 pmol的386孔板形式提供。寡核苷酸以(从水中)冻干样品形式运输 推荐的点样浓度：40 μmol/L 推荐的点样表面：聚L-赖氨酸包被的载玻片	寡核苷酸集以96孔板或者386孔板形式提供，每孔含有1400 pmol寡核苷酸，提供形式为冻存于水中或者按照客户要求的浓度和缓冲液。我们提供我们为针和环(ping & ring)系统以及使用开叉点样针(split-pin)的点样仪优化的缓冲液，或者用户也可以提供其他缓冲液供我们使用 推荐点样表面：环氧的和全部主要的表面
提供的信息	具有搜索功能的交互式的网络界面。基因清单，板及孔的位置，存取号，序列信息，寡核苷酸性质，基因注释和GO的指派，与NCBI Entrez和UniGene的链接，特有的关于mRNA剪接变体的LEADS数据，同源性报告，SAGE表达报告，以及其他更多内容 关于基因清单和有关技术方面更多的信息请参见http//www. labonweb. com/chips	基因组寡核苷酸集有配套的用户手册和光盘，包含了带有存取号和孔的位置的基因清单，T_m分布图，一个产品信息自述(readme)文件，以及样品重悬和点样的实验方案 关于Array-Ready寡核苷酸集的更多信息可以通过网络上Operon微阵列数据库(OMAD)得到。这个数据库提供了与对应的基因组数据库的链接，基因序列、染色体上位置，和每个对应的寡核苷酸的BLAST搜索结果	基因清单和有关技术方面更多的信息可以从http://www. mwg-biotech. com得到

57

a 网址：http//www. cgen. com

b 网址：http//www. operon. com

c 网址：http//www. mwg-biotech. com

d 注：Invitrogen、Genosys、Transgenomics和Illumina提供定制合成寡核苷酸集的服务

参考文献

58

Adams M., Kelly J., Gocayne J., Dubnick M., Polymeroppoulos M., Xiao H., Merril C., Wu A., Olde B., Moneno R., Kerlavage A., McCrombie R., and Venter J.C. 1991. Complementary DNA sequencing: Expressed sequence tags and Human Genome Project. *Science* **252:** 1651–1656.

Andrews J., Bouffard G.G., Cheadle C., Lu J., Becker K.G., and Oliver B. 2000. Gene discovery using computational and microarray analysis of transcription in the *Drosophila melanogaster* testis. *Genome Res.* **10:** 2030–2043.

Bhatia P., Taylor W.R., Greenberg A.H., and Wright J.A. 1994. Comparison of glyceraldehyde-3-phosphate dehydrogenase and 28S-ribosomal RNA gene expression as RNA loading controls for northern blot analysis of cell lines of varying malignant potential. *Anal. Biochem.* **216:** 223–226.

Bonaldo M.F., Lennon G., and Soares M.B. 1996. Normalization and subtraction: Two approaches to facilitate gene discovery. *Genome Res.* **6:** 791–806.

Brown P.O. and Botstein D. 1999. Exploring the new world of the genome with DNA microarrays. *Nat. Genet.* **21:** 33–37.

Cheung V.G. and Nelson S.F. 1998. Genomic mismatch scanning identifies human genomic DNA shared identical by descent. *Genomics* **47:** 1–6.

Christoffels A., van Gelder A., Greyling G., Miller R., Hide T., and Hide W. 2001. STACK: Sequence Tag Alignment and Consensus Knowledgebase. *Nucleic Acids Res.* **29:** 234–238.
Davis H.A., Wong D.T., Colbert I., Soares S., Sorge J.A., and Mullinax R.L. 2001. Normalize and validate array systems using exogenous nucleic acid controls. *Strategies Newsl.* **13(4):** 128–130 (http://www.stratagene.com/vol13_4/p128-130.htm)
DeRisi J.L., Iyer V.R., and Brown P.O. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278:** 680–686.
Diehl F., Grahlmann S., Beier M., and Hoheisel J.D. 2001. Manufacturing DNA microarrays of high spot homogeneity and reduced background signal. *Nucleic Acids Res.* **29:** E38.
Drexler H. 1988. Bacteriophage T1. In *The bacteriophages* (ed. R. Calendar), vol. 1, p. 236. Plenum Press, New York.
Eickhoff B., Korn B., Schick M., Poustka A., and van der Bosch J. 1999. Normalization of array hybridization experiments in differntial gene expression analysis. *Nucleic Acids Res.* **27:** e33.
Goffeau A., Barrell B.G., Bussey H., Davis R.W., Dujon B., Feldmann H., Galibert F., Hoheisel J.D., Jacq C., Johnston M., Louis E.J., Mewes H.W., Murakami Y., Philippsen P., Tettelin H., and Oliver S.G. 1996. Life with 6000 genes. *Science* **274:** 546–567.
Halgren R.G., Fielden M.R., Fong C.J., and Zacharewski T.R. 2001. Assessment of clone identity and sequence fidelity for 1189 IMAGE cDNA clones. *Nucleic Acids Res.* **29:** 582–588.
Hegde P., Qi R., Abernathy K., Gay C., Dharap S., Gaspard R., Hughes J.E., Snesrud E., Lee N., and Quackenbush J. 2000. A concise guide to cDNA microarray analysis. *BioTechniques* **29:** 548–556.
Hughes T.R., Mao M., Jones A.R., Burchard J., Marton M.J., Shannon K.W., Lefkowitz S.M., Ziman M., Schelter J.M., Meyer M.R., Kobayashi S., Davis C., Dai H., He Y.D., Stephaniants S.B., Cavet G., Walker W.L., West A., Coffey E., Shoemaker D.D., Stoughton R., Blanchard A.P., Friend S.H., and Linsley P.S. 2001. Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. *Nat. Biotechnol.* **19:** 342–347.
Iyer V.R., Horak C.E., Scafe C.S., Botstein D., Snyder M., and Brown P.O. 2001. Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature* **409:** 533–538.
Iyer V.R., Eisen M.B., Ross D.T., Schuler G., Moore T., Lee J.C.F., Trent J.M., Staudt L.M., Hudson J., Jr., Boguski M.S., Lashkari D., Shalon D., Botstein D., and Brown P.O. 1999. The transcriptional program in the response of human fibroblasts to serum (see comments). *Science* **283:** 83–87.
Killmann H., Benz R., and Braun V. 1996. Properties of the FhuA channel in the *Escherichia coli* outer membrane after deletion of FhuA portions within and outside the predicted gating loop. *J. Bacteriol.* **178:** 6913–6920.
Killmann H., Videnov G., Jung G., Schwarz H., and Braun V. 1995. Identification of receptor binding sites by competitive peptide mapping: Phages T1, T5, and ϕ80 and colicin M bind to the gating loop of FhuA. *J. Bacteriol.* **177:** 694–698.
Knight J. 2001. When the chips are down. *Nature* **410:** 860–861.
Lennon G., Auffray C., Polymeropoulos M., and Soares M.B. 1996. The I.M.A.G.E. Consortium: An integrated molecular analysis of genomes and their expression. *Genomics* **33:** 151–152.
Mansur N.R., Meyer-Siegler K., Wurzer J.C., and Sirover M.A. 1993. Cell cycle regulation of the glyceraldehyde-3-phosphate dehydrogenase/uracil DNA glycosylase gene in normal human cells. *Nucleic Acids Res.* **21:** 993–998.
Miki R., Kadota K., Bono H., Tomaru Y., Carninci P., Itoh M., Shibata K., Kawai J., Konno H., Watanabe S., Sato K., Tokusumi Y., Kikuchi N., Ishii Y., Hamaguchi Y., Nishizuka I., Goto H., Nitanda H., Satomi S., Yoshiki A., Kusakabe M., DeRisi J.L., Eisen M.B., Iyer V.R., Brown P.O., Muramatsu M., Shimada H., Okazaki Y., and Hayashizaki Y. 2001. Delineating developmental and metabolic pathways in vivo by expression profiling using the RIKEN set of 18,816 full-length enriched mouse cDNA arrays. *Proc. Natl. Acad. Sci.* **98:** 2199–2204.
Miller R.T., Christoffels A.G., Gopalakrishnan C., Burke J., Ptitsyn A.A., Broveak T.R., and Hide W.A. 1999. A comprehensive approach to clustering of expressed human gene sequence: The sequence tag alignment and consensus knowledge base. *Genome Res.* **9:** 1143–1155.
Nature Genetics. 1999. The chipping forecast (supplement), vol. 21, no. 1. Nature America Inc., New York, New York.
Phimister B. 1999. Going global. *Nat. Genet.* **21:** 1.
Pluthero F.G. 1993. Rapid purification of high-activity Taq DNA polymerase. *Nucleic Acids Res.* **21:** 4850–4851.

Reid J.L., Iyer V.R., Brown P.O., and Struhl K. 2000. Coordinate regulation of yeast ribosomal protein genes is associated with targeted recruitment of Esa1 histone acetylase. *Mol. Cell* **6:** 1297–1307.

Ren B., Robert F., Wyrick J.J., Aparicio O., Jennings E.G., Simon I., Zeitlinger J., Schreiber J., Hannett N., Kanin E., Volkert T.L., Wilson C.J., Bell S.P., and Young R.A. 2000. Genome-wide location and function of DNA binding proteins. *Science* **290:** 2306–2309.

Sambrook J. and Russell D.W. 2001. *Molecular cloning: A laboratory manual*, 3rd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Spanakis E. 1993. Problems related to the interpretation of autoradiographic data on gene expression using common constitutive transcripts as controls. *Nucleic Acids Res.* **21:** 3809–3819.

Ten Bosch J., Seidel C., Batra S., Lam H., Tuason N., Saljoughi S., Zhou H., and Saul R. 2001. Operon® Array-Ready Oligo Sets™ provide sequence-optimized 70 mers for DNA microarrays. *QIAGEN News* issue no. 4 (http://www.qiagen.com/literature/qiagennews).

Velculescu V.E., Zhang L., Zhou W., Vogelstein J., Basrai M.A., Bassett Jr., D.E., Hieter P., Vogelstein B., and Kinzler K.W. 1997. Characterization of the yeast transcriptome. *Cell* **88:** 243–251.

Wei Y., Lee J.M., Richmond C., Blattner F.R., Rafalski J.A., and LaRossa R.A. 2001. High-density microarray-mediated gene expression profiling of *Escherichia coli. J. Bacteriol.* **183:** 545–556.

Yang Y.H., Dudoit S., Luu D., Lin D.M., Peng V., Ngai J., and Speed T.P. 2002. Normalization for cDNA microarray data: A robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res.* **30:** e15.

Yue H., Eastman P.S., Wang B.B., Minor J., Doctolero M.H., Nuttall R.L., Stack R., Becker J.W., Montgomery J.R., Vainer M., and Johnston R. 2001. An evaluation of the performance of cDNA microarrays for detecting changes in global mRNA expression. *Nucleic Acids Res.* **29:** E41–41.

因特网资源

59

http://cgap.nci.nih.gov The Cancer Genome Anatomy Project (CGAP)

http://genome.rtc.riken.go.jp/home.html RIKEN Genome Exploration Research Group, Genome Science Laboratory

http://genome.wustl.edu/est/mouse_esthmpg.html Washington University Mouse EST Project

http://image.llnl.gov I.M.A.G.E. Consortium (Integrated Molecular Analysis of Genomes and their Expression

http://image.llnl.gov/image/html/idistributors.shtml The I.M.A.G.E. Consortium Distributors

http://image.llnl.gov/image/imagene/current/bin/search IMAGEne Database

http://mgc.nci.nih.gov Mammalian Gene Collection (MGC), National Cancer Institute

http://mwgbiotech.com MWG Biotech Homepage

http://www.allgenes.org All Genes database. Computational Biology and Informatics Laboratory at the University of Pennsylvania.

http://www.apbiotech.com/lsn–home/lsn–index.html Samartzidou H., Turner L., and Houts T. 2001. Lucidea™ Microarray ScoreCard: An integrated tool for validation of microarray gene expression experiments. Amersham Biosciences LSN On-line

http://www.arrayit.com TeleChem International, Inc.

http://www.atcc.org American Type Culture Collection (ATCC)

http://www.atcc.org/SearchCatalogs/tasc2.cfm#who ATCC: Who makes ESTs

http://www.beckman.com Beckman Coulter, Inc. Homepage

http://www.biorobotics.co.uk/Pages/replicat.html BioRobotics Replication Accessories

http://www.cbc.umn.edu/ResearchProjects/Arabidopsis/index.html *Arabidopsis* cDNA Sequence Analysis Project

http://www.cbil.upenn.edu Computational Biology of Informatics Laboratory at the University of Pennsylvania.

http://www.cgen.com Compugen Ltd.Homepage

http://www.fruitfly.org Berkeley *Drosophila* Genome Project (BDGP)

http://www.fruitfly.org/sequence/genomic-clones.html BDGP Release 2.5 Drosophila Genomic Sequence

http://www.genetix.com Genetix Homepage

http://www.genome.washington.edu/uwgc/analysistools/repeatmask.htm RepeatMasker (University of Washington Genome Center

http://www.jgi.doe.gov/programs/instrumentation/preptrack.htm JGI Programs, DOE Joint Genome Institute operated by The University of California for the U.S. Department of Energy Lawrence Berkeley National Lab Human Genome Center (Kaaren Q'd David-see Liq.Handling)

http://www.labonweb.com/chips LabOnWeb Oligolibraries
http://www.lionbioscience.com LION Bioscience
http://www.mwg-biotech.com MWG Biotech AG
http://www.ncbi.nlm.nih.gov/BLAST National Center for Biotechnology Information, BLAST Info
http://www.ncbi.nlm.nih.gov/dbEST/index.html Expressed Sequence Tags Database
http://www.ncbi.nlm.nih.gov/gorf/gorf.html Open Reading Frame (ORF) Finder
http://www.ncbi.nlm.nih.gov/UniGene/index.html UniGene Resources
http://www.operon.com Operon Homepage
http://www.packardbioscience.com Packard BioScience Company Homepage
http://www.premierbiosoft.com/dnamicroarray/dnamicroarray.html Premier Biosoft International
http://www.qiagen.com QIAGEN Homepage
http://www.resgen.com ResGen: An Invitrogen Corporation
http://www.robsci.com Robbins Scientific Homepage
http://www.rzpd.de Resource Center/Primary Database (RZPD)
http://www.sanbi.ac.za/Dbases.html South African National Bioinformatics Institute (SNABI)
http://www.stat.Berkeley.EDU/users/terry/zarray/Html/normspie.html) Yang Y.H., Dudoit S., Luu P., and Speed T. 2001. Normalization for cDNA microarray data. Speed Group Microarray Page.
http://www.stratagene.com/vol13_4/p128-130.htm Stratagene Inc. *Strategies Newsletter*.
http://www.tecan-us.com TECAN U.S.
http://www.tigr.org The Institute for Genomic Research (TIGR)
http://www.tigr.org/tdb/egad/egad.shtml The Expressed Gene Anatomy Database (EGAD)
http://www.tigr.org/tdb/hgi/index.html TIGR Human Gene Index
http://www.tigr.org/tdb/index.shtml TIGR Databases

第 2 章　玻璃微阵列的点样

导言

在固相支持物表面将 DNA 样品按有规则的间距进行排列，这在概念上比较直观，但要想得到一个精确和可靠的微阵列制备方法却具有一定的挑战性。本部分主要讨论用于表达分析的玻片微阵列的制备。有关膜微阵列的制备与应用见第 4 章。

玻片微阵列首先产生于斯坦福大学（Schena et al. 1995）。微阵列由一台 *xyz* 轴的自动移动平台装置（gantry robot）制备，通过装置上成排的点样针，将 96 孔板中的 DNA 溶液点样到一系列经过预先处理的玻片表面。尽管在不久的将来，它有可能被其他方法所取代，但到本书写作时为止，作为“内部”制备微阵列的一种手段，该方法仍然被广泛应用。最初，商业化的微阵列制备装置非常有限，且这一技术的早期开拓者们均使用了由 Pat Brown 小组提供的详细说明书来制造他们自己的点样仪（http://cmgm.stanford.edu/pbrown/mguide/index.html）。自制点样仪仍然是一种有效且经济的选择。然而，商品化的自动设备越来越多，可以将打印头（printing head）精确地定位于一大片玻璃载玻片的上方，许多研究者选择购买这些装置作为进入微阵列领域的一种快速的、相对省力的（但不便宜）途径。在表 2-1 及本章后关于**微阵列接触印刷使用的点样仪与点样针**信息栏中，总结了一些这样的仪器和影响它们性能的因素。

方案 2 中描述了利用接触式点样仪对玻片微阵列进行点样的常规步骤。该方案由**微阵列接触印刷使用的点样仪与点样针**信息栏支持，在信息栏中讨论了决定制备的微阵列质量的一些因素。其中的关键因素包括点样针（printing pin）[也可称为针头（tip）、

表 2-1　点样仪：

公司	印刷全平台 10K 玻片需要的时间	玻片容量/固定方式	尺寸 /cm^3 (W×D×H)	清洗和超声处理	点样针
Affymetrix：affymetrix. com					
Affymetrix 417 Arrayer	25 h（包括清洗与干燥时间）	42/弹簧夹和对准杆	84×53×49	流动液体，真空干燥	1 或 4 针和环系统，装样量 1.5 μl 50 pl 点样量
Affymetrix 427 Arrayer	12.5 h（包括清洗与干燥时间）	42/弹簧夹和对准杆	84×53×49	流动液体，文氏管 Venturi 系统（压缩空气）	8 针和环系统，装样量 1.5 μl 50 pl 点样量
BioRobotics：www. BioRobotics. com					
Microgrid 2	8 h 完成 108 片载玻片	108	90×75×76	真空/清洗	实心针（BioRobotics）或者鹅毛笔针（Telechem） 1～64
Microgrid 2 Compact	4 h	48	60×70×74	真空/清洗	实心针（BioRobotics）或者鹅毛笔针（Telechem） 1～48
Microgrid 1	8 h 完成 94 片载玻片	84		清洗	实心针（BioRobotics）或者鹅毛笔针（Telechem） 1～48
Cartesian：www. cartesian. com					
MicroSys 5100	3.5 h	10	60.5×35.8×38	真空/清洗	Quill Telechem 32
PixSys 5500	6 h	50	76.2×81.22×65.5	真空/清洗	Quill Telechem 32 或 48
ProSys 5510	6.5 h	100	71×74×36.5	真空/清洗/超声	Quill Telechem 32 或 48

规格

间距，样点尺寸，最大密度	分辨率单位/μm	堆板器加盖和取盖的容量多孔板的格式	湿度	空气过滤	其他特性
用户自定义，10 μm 间隔 150～200 μm >30 000	精度：±10 重复性：±2	无 3 块多孔板槽，手工进板	无	无	点样一致性（自始至终不发生样点缺失）；集成的系统：点样，扫描，跟踪软件，分析软件，数据挖掘软件；致力于现场服务和现场支持
用户自定义，10 μm 间隔 150～200 μm >30 000	精度：±10 重复性：±2	无 3 块多孔板槽，手工进板	有	无	点样一致性（自始至终不发生样点缺失）；集成的系统：点样，扫描，跟踪软件，分析软件，数据挖掘软件；致力于现场服务和现场支持
10 μm	1.5	集成堆板器；为 24 块多孔板加盖 96/384/1536	有/extra	有/extra	克隆追踪 Softtouch 点样技术 条码阅读 多孔板冷却 点样膜印刷
10 μm	1.5	容量 1～4 块多孔板	有/extra	有/extra	克隆追踪 Softtouch 点样技术 条码阅读
200 μm	5	集成堆板器；为 24 块多孔板加盖 96/384	无	有/extra	克隆追踪 条码阅读 点样膜印刷
150 μm	1.5	无	有	—	无
150 μm	1.5	可选择	有	—	克隆追踪
150 μm	1.5	可选择	有	—	克隆追踪，条码，显示系统，超声清洗工作站

公司	印刷全平台10K玻片需要的时间	玻片容量/固定方式	尺寸/cm^3 (W×D×H)	清洗和超声处理	点样针
GeneMachine：www. genemachines. com					
OminGrid	100片玻片点样10 000点要3.5 h；使用48针点样，一块384孔板可以在少于7.5 min内点样到100片载玻片上；在小于8.5 h 内，28 000样点	100	135×143×153	可定制的组合，包括超声，水/使用蠕动泵的流体清洗/和真空干燥	1～48 Telechem quill或实心针，Majer Precision quill；点样针可从GeneMachines得到
OminGrid Accent	50片玻片点样10 000点要2.5 h；使用48针点样，1 h内可以按照每片4000点点样50片载玻片	50	104.5 × 74 ×53.3 130×130×160	可定制的组合，包括超声，水/使用蠕动泵的流体清洗/和真空干燥	1～48 Telechem quill或实心针，Majer Precision quill；点样针可从GeneMachines得到
Genetix：www. genetix. com					
QArray	使用24根Genetix针尖直径150 μm实心微阵列点样针，26 h可印刷84×10K载玻片；24根multispot（分叉）点样针要8.5 h；时间包括针的装载，清洗/干燥循环（可变的）	84/弹簧夹位于6个可移动的载玻片固定器中	143×73×175 包括工作台面	高压下3种不同的清洗试剂和压缩空气干燥	最多24根Genetix针尖直径150 μm实心微阵列点样针；经过认证可使用其他制造商的点样针；可选择最多48针

续表

间距，样点尺寸，最大密度	分辨率 单位/μm	堆板器 加盖和取盖的容量 多孔板的格式	湿度	空气过滤	其他特性
用户自定义，2.5 μm 间隔，最小 10 μm 100～200 μm（取决于针的形状）每张载玻片 10 000＋样点	精度：小于 ±2.5 位置分辨率：2.5 μm（x 和 y 轴）；1.25 μm（z 轴）	可选的伺服手臂用于处理 72 块 96 或 384 孔板；长时间，无人值守的工作；移去多孔板盖的性能；可以随机访问加样板	有 通过传感器和控制器可控制最高为 80%	有 HEPA 过滤用于无尘印刷	标准的密封罩；可控环境下清楚的操作台。灵活的软件，多重复制选项；校准设置，多孔板载入，用户定义的点样区域；基底，和洗涤方案；样品追踪特性将 OminiGrid 的阵列数据与扫描仪和分析软件集成。 定制的基底尺寸和类型；膜点样；吸墨垫（blotting pad） 现场的和工厂中的应用支持 可选件：条码系统 包括了 Incyte 的许可
用户自定义，2.5 μm 间隔，最小 10μm 100～200 μm（取决于针的形状）每张载玻片 10 000＋样点	2.5	无 3 块 96 孔或 384 孔板	可选的 通过传感器和控制器可控制最高为 80%	有 HEPA 过滤用于无尘印刷	自带笔记本电脑；多重阵列设计软件，多孔板校准；多孔板载入，用户定义的点样区域；基底，和洗涤方案；样品追踪特性将 OminiGrid 的阵列数据与主要的扫描仪和分析软件集成。 定制的基底尺寸和类型；膜点样；吸墨垫（blotting pad） 全面的现场和工厂中支持 包括了 Incyte 的许可
使用 Genetix 150 μm 点样针，样点直径 250 μm，间隔 280 μm；间隔取决于阵列的样式，用户可选，步长 1 μm	所有轴（x、y 和 z）均为 1 μm	可选的堆板器，最多 70 块 Genetix low-profile 96 孔或 384 孔板，带有自动取盖和加盖功能；没有堆板器时板的容量为 5 块 96 孔或 384 孔板	有 标准的	有 标准的	灵活的样式设定，点样针的数量可变，如 1、2、4、8、16 和 24；样点间距可变，步长 1 μm；可选择的多根点样针的校准和调节点样开始时针尖在多孔板中的深度；连接的数据库（QSoft Library Manger）以保存多孔板的信息来生成输出的数据文件，用于数据分析时的样点追踪；软件可以追踪加样板和载玻片的条码；可选的堆板器，能够自动打开加样板盖

公司	印刷全平台10K玻片需要的时间	玻片容量/固定方式	尺寸/cm^3 (W×D×H)	清洗和超声处理	点样针
GeneMachine：www. genemachines. com					
QArray^{lite}	数据缺失	90/真空吸持 6行独立的载玻片，每行15片	143×73×175 仅仅是QArray Lite单元	高压下2种不同的清洗试剂和压缩空气干燥；可选项包括3个瓶子，1个瓶子，或者超声洗浴	最多24根Genetix针尖直径150 μm实心微阵列点样针；经过认证可使用其他制造商的点样针；可选择最多48针
QArray^{mini}	数据缺失	54/真空吸持	93×65×52	高压下2种不同的清洗试剂和压缩空气干燥	最多24根Genetix针尖直径150 μm实心微阵列点样针
Intelligent Bioinstrument：www. intelligentbio. com					
IBI Microarrayer HT	用户可调	100	75×25	自动超声清洗系统，多针真空干燥工作站	12～48根点样针
Packard BioScience：www. packardbioscience. com					
SpotArray Enterprise	300样品/h，到108片载玻片上；循环时间取决于希望的清洁度和点印的材料	108/真空；可以使用定制的芯片和微孔板（非载玻片）	173.4 × 122.4 ×130	2个清洗工作站；定制的内部或者外部超声清洗	8根压电控制针尖(piezo tip)
SpotArray 24	1536 (4×384) 20 min内在20片载玻片上复制样品	24/定位的使用高精度的记录执行器（registration actuator）	仪器模块：66.3×64×41 60 lbs磅	4束加压的射流/针，真空干燥	1～4 Stealth点样针

续表

间距，样点尺寸，最大密度	分辨率 单位/μm	堆板器 加盖和取盖的容量 多孔板的格式	湿度	空气过滤	其他特性
使用 Genetix 150 μm 点样针，样点直径 250 μm，间隔 280 μm；间隔取决于阵列的样式，用户可选，步长 1 μm。	所有轴（x、y 和 z）均为 1 μm。	没有堆板器；板容量为 5 块 96 或者 384 孔板	可选的	可选的	灵活的样式设定，点样针的数量可变，如 1、2、4、8、16 和 24；样点间距可变，步长 1 μm；可选择的多根点样针的校准和调节点样开始时针尖在多孔板中的深度；连接的数据库（QSoft Library Manger）以保存多孔板的信息来生成输出的数据文件，用于数据分析时的样点追踪；加样板固定器的高度可调节，能够使用一系列类型的多孔板
使用 Genetix 150 μm 点样针，样点直径 250 μm，间隔 280 μm；间隔取决于阵列的样式，用户可选	1 μm（x 轴），1 μm（y 轴），50 μm（z 轴）	没有堆板器；板容量为 5 块 96 或者 384 孔板	可选的	可选的	灵活的样式设定，点样针的数量可变，如 1、2、4、8 和 16；QSoft Library Manger 用来生成输出的数据文件，用于数据分析时的样点追踪；加样板固定器的高度可调节，能够使用一系列类型的多孔板
5000 样点/cm²，20 000 样点/载玻片，100～150 μm，75～100 μm 样点的点样针可供选择	0.55 μm（x 轴），0.88 μm（y 轴），1.63 μm（z 轴）	自动的多孔板和板盖操作；最多 108 块多孔板；96/384 孔板	可选的	可选的	多孔板冷冻储藏室 可编程的样点间隔 吸墨工作站 条码阅读器
1600 样点/cm²；尺寸：150～200 μm，取决于基底。	0.1 μm 精度；8 μm 准确度，1 μm 精度。	最多 50 块板，冷却到 4 ℃ 96/384 孔板	压电打印不可使用	无	非接触式分配；可以处理定制尺寸的芯片；灵活的样品选择（DNA，蛋白质，细胞，小珠，肽）；样点大小一致，变异系数小于 6%
最多 81 920 样点/载玻片，直径 75～150 μm	重复性：±10	1～4 块 96 或 384 孔板；取盖器	控制在 50%～65%RH	0.3 μm HEPA 过滤器	点样针提升装置生成更加紧密的阵列；多孔板条码阅读器；取盖器和传感器确保板盖已经移去；可选的平板监视器

公司	印刷全平台10K玻片需要的时间	玻片容量/固定方式	尺寸/cm^3（W×D×H）	清洗和超声处理	点样针
SpotArray 72（见SpotArray 24）	1 h内在68片载玻片上复制样品	72片载玻片	100×70×43 130 lbs磅	4束加压的射流/针，真空干燥	1～4 Stealth点样针
Biochip Arrayer	1 h内将100个样品点样到20片载玻片上	25/真空吸持；可以使用定制的芯片和微孔板	135×84×58.6	2个清洗工作站；定制的内部或者外部超声清洗	4根压电控制针尖
TeleChem：www. arrayit. com					
SpotBot Personal Microarrayer	30 h（包括进样和清洗/干燥时间）	14/弹簧带准直杆	30×30×24	流体浴，压缩空气管干燥；可编程的	1、2或4根专利的Stealth Microspotting Pin；每次取样0.25 μl；点样量0.5 nl；20种打印头设计
Virtek Vision Corp：www. virtekbiotech. com					
Chipwriter Pro	约8 h	126片玻璃载玻片或者15张点样膜	127×97×126	真空；流过式清洗工作站；超声清洗器	最多48根TeleChem Stealth quill pin或者48根TeleChem实心针
Chipwriter Compact	约6 h	24片玻璃载玻片或者2张点样膜	60×87.5×67.5	真空；2个清洗工作站；超声清洗器；灵活的，用户自定义的点样针清洁方案	最多48根TeleChem Stealth quill pin或者48根TeleChem实心针

续表

间距，样点尺寸，最大密度	分辨率 单位/μm	堆板器 加盖和取盖的容量 多孔板的格式	湿度	空气过滤	其他特性
最多 81 920 样点/载玻片，直径 75～150 μm	重复性：±10	1～4 块 96 或 384 孔板；取盖器	控制在 50%～65%RH	0.3 μm HEPA 过滤器	点样针提升装置生成更加紧密的阵列；多孔板条码阅读器；取盖器和传感器确保板盖已经移去；可选的平板监视器与 SpotArrayer 24 相比，在双倍时间内，样点多出 3 倍
1600 样点/cm^2；尺寸：150～200 μm，取决于基底。	分辨率：2 μm 准确度：20 μm 重复性：±10	最多 4 块加样板，冷却到 4 ℃；	压电打印不可使用	去除静电的离子发生器	非接触式分配；可以处理定制尺寸的芯片；灵活的样品选择(DNA、蛋白质、细胞、小珠、肽)；样点大小一致，变异系数小于 5%；用于观察微结构的目标照相机
用户自定义，1 μm 步长；90～300 μm 取决于使用的 Stealth Pin；约 50 000	精度：±10 重复性：±10	无 1 384 孔板 手工装入	即将提供	即将提供	配有 Mini-Stealth Micro Spotting 装置（美国专利号 6101946)；市场上最便宜的点样仪；送货时已经校准并包括作图软件；方便的自我安装/设置；完全便携；可点样多种样品类型，包括 DNA、蛋白质、糖类、脂类和更多
用户自定义，1.25 μm 步长，最小中心-中心：120 μm 样点直径： Microspotting Pin 75～400 μm; 实心针：75～500 μm; 最大密度： 123 333 样点/载玻片	1.25 μm（x 和 y 轴)； 0.25 μm（z 轴）	有 板容量为 18 或者 32 块 96 或 384 孔板	有	有	克隆追踪；液体操作；克隆选取；膜阵列；最多点印 250 样点/从加样板取样；固定的平台；适合大多数基因组和蛋白质组应用
用户自定义，2.5 μm 步长，最小中心-中心：150 μm 样点直径： Microspotting Pin 75～400 μm; 实心针：75～500 μm; 最大密度： 78 933 样点/载玻片	1.22 μm（x 和 y 轴)； 0.61 μm（z 轴）	没有堆板器 板容量为 1 块 96 或 384 孔板	有	有	内建样品追踪；可以点印载玻片与点样膜 固定的平台；适合大多数基因组和蛋白质组应用

笔（pen）或鹅毛笔（quill）］的形状、重复性及耐久性。不平整的针在点样循环中所输送DNA溶液的量也不一样，这就增加了后面图像分析程序的负担。当玻片经过杂交和扫描以后，由针带来的问题将变得非常明显，第3章的疑难问题解答指南中对此进行了详细描述。一些点样仪中提供了合适的点样针，而其他点样仪则是使用专业制造商提供的针。方案3对针的保养与清洗过程进行了叙述，同时介绍了商品化点样针的获取来源。

63

另一个影响样点形状的重要因素是用来重悬DNA的溶液。关于不同点样溶液的讨论，参考第1章的方案5。第1章的方案2和方案3介绍了制备现成的点样材料。点样材料的特性与载玻片基底的选择是紧密相连的。**用于点样的基底**信息栏中提供了一些常用基底的详细信息及可用性。方案1中描述了多聚L-赖氨酸包被的载玻片的制备以及点样前的处理。

由于皮升量级液点递送的物理机制，样点质量及阵列的灵敏度最终取决于印刷头的运动、点样针的性能和DNA溶液的物理及化学特性，以及点样表面之间的相互作用。考虑到这些因素，我们除了系统地确定每个研究设施的特定平台所适合的这些因素的最佳组合之外，别无他法。

68

方案1

制备用于微阵列的聚L-赖氨酸玻片

DeRisi 实验室（加利福尼亚州立大学，旧金山，加利福尼亚，94143）

多聚L-赖氨酸是用于微阵列分析中最有效的玻片包被试剂之一；有关基底、表面及包被试剂的讨论，请见本章后关于用于点样的基底的信息栏。本方案描述了试剂的制备以及用多聚L-赖氨酸对每批180张玻片进行包被的方法。要注意，DeRisi实验室不推荐使用商业化已包被的玻片。

- **确保玻片架中间部分被轻轻的向内弯曲，以将玻片固定得更加安全！**
- **使用丁腈手套，在任何时候都不要用带粉的手套！**
- **为了防止灰尘，时刻将玻片盖住或者浸没在溶液中！**
- **使用不含 Ca^{2+}、Mg^{2+} 的磷酸缓冲盐溶液（PBS）！**
- **多聚L-赖氨酸的量取只能使用塑料的实验室用具！**

警告：戴好防护眼镜、手套，穿好实验服！

材料

注意：标记有〈!〉的材料的正确处理方法见附录3。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成见附录2。

使用前将储存液稀释到合适的浓度。

乙醇（95%和70%）

据报道，有些批次的乙醇（100%或200标准强度）中含有会使微阵列试验中产生很深荧光背景的污染物质，95%的乙醇不存在这个问题。为避免这个问题，可以用95%的乙醇或者用已证明不会产生高荧光背景的无水乙醇。Warner-Graham公司（Cockeysville，Maryland）出售的无水乙醇（USP级乙醇，64-17-5）效果很好。

H_2O（灭菌）

NaOH颗粒〈!〉

1×磷酸盐缓冲液（PBS）（已过滤），不含Mg^{2+}和Ca^{2+}

Poly-L-lysine溶液（0.1% m/V）（P 8900，Sigma）

专用装置

69

水平摇床

显微镜载玻片（玻璃，Gold Seal；12-518-100A，Fisher）

载玻片盒（塑料）

载玻片容器（玻璃；ThermoShandon）

载玻片架（金属；ThermoShandon）

配有多孔板适配器（plate carrier）的台式离心机（如GS-6KR Allegra 6R系列，Beckman Instruments）

真空烘箱（50℃）

方法

1. 准备载玻片清洗液
   a. 将200 g的NaOH颗粒溶解于800 ml已灭菌的H_2O。
   注意：这是一个强放热反应！戴好防护眼镜、手套，穿好实验服！
   b. 向溶液中加入额外的NaOH颗粒直至饱和。
   c. 向NaOH溶液加入1200 ml 95%的乙醇。
   d. 向溶液中加入额外的H_2O直至它变澄清。
2. 用H_2O彻底漂洗载玻片盘。在每个载玻片盘内放置一个排列了Gold Seal显微镜载玻片的载玻片架（30片/架）。
3. 用步骤1中准备的清洗液浸没载玻片，将载玻片盘放置在振动平台上，轻摇载玻片盘至少1～6 h。

   此处清洗可除去“预先清洁载玻片”上残留的油脂和碎片。
4. 每架载玻片单独用蒸馏水清洗约30 s，确保每片玻片的正反两面都清洗过。清洗完成后，将玻片架浸入到装有水的4 L大口烧杯中。将架子堆叠好，以便它们都能放入同一个烧杯中。最好将载玻片浸在水中过夜，以确保除去全部的NaOH。

   可供选择的办法：将玻片架浸入装有水的载玻片盘里，清洗30 min，接着转移到另一个

装有水的载玻片盘中。重复该清洗过程 4 次。

5. 为每个载玻片盘中制备多聚 L-赖氨酸溶液，在摇动的平板上混合下列试剂：

H_2O	285 ml
PBS（过滤的），不含 Ca^{2+} 和 Mg^{2+}	35 ml
多聚 L-赖氨酸溶液	30 ml

6. 将每个架子从装有水的 4 L 大口烧杯中取出，让过量的液体沥干。再将载玻片浸没到多聚 L-赖氨酸溶液中，然后在水平摇床上轻轻搅动溶液中的载玻片30～60 min。

7. 分别漂洗每个架子上的载玻片，依次从多聚 L-赖氨酸溶液中取出每架载玻片，等其短暂的沥干后，浸入到装有新鲜水的 4 L 大口烧杯中。轻轻地前后转动载玻片架子进行漂洗，然后把它放到台式离心机（GS-6KR 或者类似的）中。

70

如果有数批载玻片需要处理，可以重复使用多聚 L-赖氨酸溶液。每次使用过后，溶液经过简单的过滤，可以在 4℃下保存在塑料容器中。过滤的溶液最多使用 4 次，但必须在 2 个月后丢弃。

8. 将玻片架在最低速度下（540 r/min）离心 2 min，以干燥玻片。

玻片架放入之前，首先将离心机内部擦干净，防止可能转移到玻片上的灰尘。

9. 将玻片放入 50℃的真空烘箱里干燥 10～30 min。

10. 在开始 DNA 点样之前，将玻片在清洁的塑料盒中存放至少 14 天。

注意事项：我们不建议使用处理时间超过 4 个月的玻片。玻片疏水性的测试可以通过观察一滴水在其表面的滑动来进行。玻片表面不应该有水珠存在。

注意事项：在玻片点样之前

- 检查聚 L-赖氨酸涂层是透明的。
- 取一组样品，通过点样、杂交、扫描来检查每批玻片的质量（见方案 4）。

71

方案 2

在玻片上点样

Aldo Massimi，**Thomas Harris**，**Geoffrey Childs**（阿尔伯特·爱因斯坦医学院，布朗克斯，纽约，10461）；

Shauna Somerville（华盛顿卡内基研究院，斯坦福，加利福尼亚，94305）

将点样材料在玻璃载玻片上点印成密集阵列的方案随着特定点样仪的性能及组成的不同而改变（有关各种微阵列点样设备的讨论，见本章后面关于**微阵列接触印刷使用的点样仪及点样针**信息栏）。本方案提供了关于点样过程的设定及处理印刷的玻片准备用于杂交反应所必需的一般背景信息。这里对点样过程本身不作详细的描述，最好从点样仪生产厂家获取。Yue（2001）对影响玻片微阵列性能的因素进行了非常有用的分析。在开始任何点样操作之前，应该考虑下列关键因素。

阵列材料

- 利用琼脂糖凝胶检查扩增的 DNA 样品，确保只有单个条带出现。纯化扩增的 DNA 样品，除去其中含有的引物、盐、核酸和蛋白质污染物（见第 1 章，方案 3）。
- 将扩增的 cDNA 样品在点样液中以 150～250 μg/ml 浓度重悬（见第 1 章，方案 5）。如果 DNA 样品是上次点样结束时干燥保存的（这是储存它的最好办法），可将它重悬数小时或者过夜（第 1 章，方案 5）。重悬后，将多孔板短暂离心，确保 DNA 溶液沉积到板孔底部。调节点样仪的印刷头，确保点样针进入孔中的深度约为 1 mm。这个预防措施用于减少吸附在点样针外侧点样溶液的量。点样针外侧的多余溶液被不均匀的分布在最初印刷的几个点中，增多了大的不规则样点。
- 将适合于要使用玻片数量的点样材料溶解在点样缓冲液中。不能使用含有游离氨基的缓冲液（如 Tris），这会抑制 DNA 结合到硅烷化或者甲基硅烷化的玻片上（见关于**用于点样的基底**信息栏）。
- 重悬用于点样的长链寡核苷酸，浓度约为 40 pmol/L。
- 设定合适的点样规则，考虑应该包含哪些对照，这些对照应该放在阵列的哪个部位，每个阵列上对照要点样多少次等。建议在每个阵列中子阵列的四个角落上包含已知会给出明亮样点的对照 DNA 样品，这样做可以使各种软件程序容易建立与印刷阵列图像对应的网格。关于进一步的考虑，见第一部分关于**点样微阵列中使用的对照**信息栏。

72

点样仪与点样针

- 在立体显微镜下检查点样针，确保外表面及凹槽（鹅毛笔针）保持清洁，并且针的末端是完整和笔直的。必要时对出现的问题进行矫正。有关针的清洁方法，见方案 4 和关于**微阵列接触印刷使用的点样仪及点样针**信息栏。
- 对点样仪进行预热，确保其能正常工作。检查湿度是否在容许限度之内，一般建议的相对湿度范围为 50%～60%。

点样仪的设置

- 如果不了解如何设定，请勿触动点样仪——点样针、平台和载玻片之间的对准必须正确无误。点样针与未能正确对准的微量滴定板表面发生摩擦将造成极大的损失。
- 使用点样仪对珍贵样品点样之前，可以用一块测试板进行试点样，确保点样仪工作正常。确保湿度（50%～60%）与温度（20℃～24℃）处在正常范围之内。检查每根点样针是否正常工作，对不正常的针进行清洁处理（方案 4）。
- 确保清洗工作站正常工作。
- 有关性能参数的讨论，包括停留时间、接触速度，见关于**微阵列接触印刷使用的点样仪及点样针**信息栏。

载玻片

- 保证玻片的高质量，其表面性质要与点样缓冲液的特性相匹配（见关于**用于点样的基底**信息栏）。
- 一些载玻片只有单面被修饰（如 ArrayIT Super Amine 玻片，http://arrayit.com/Products/Substrates/substrates.html），确保包被的一面朝上放置。
- 用玻璃刀标志载玻片；不能使用墨水。注意玻片在点样仪上的方向，保证不同点样批次间玻片上第一个样点的位置是一致的。在点样过程中不能让玻片发生移动。
- 确保有足量的载玻片库存（注意包被有多聚 L-赖氨酸的玻片储存时间不能超过 4 个月）。使用新批次载玻片进行大规模的点样之前，最好进行测试点样与杂交。
- 检查印刷头是否仍然与平台上所有的载玻片对准。载玻片尺寸的微小误差，通过平台上全部载玻片的累加，将会导致载玻片位置较大的错位（图 2-1）。



73

图 2-1　点样平台上载玻片的布局图

（左）尺寸规则的载玻片在点样平台上正确对准；（右）尺寸不规则的载玻片不能正确对准

点样过程

为实现执行良好的点样操作，应当：

- 确保在整个点样过程中，密切监视仪器。如果点样过程时间较长，应为参加的每个人草拟一张明确的计划表，采取两人一组的办法。
- 将玻片排列在平台上，并将其固定。
- 再次检查点样仪的设置。
- 一旦点样过程开始后，要确保加样板的更换是按照计划表上进行的。
- 可以偶尔中止点样过程，用放大镜检查玻片上的样点，看点样针是否堵塞以及样点是否过大。若所有样点的无规律性，则表明湿度、点样液体积或点样仪的设置可能出现问题。如果是个别的样点出现问题，则表明对应的点样针出现问题，应该取下进行仔细检查。
- 如果偶尔有载玻片上出现样点缺失，那么 z 轴设置可能出现问题，不平整的工作台面会加剧这一问题。

- 点样完成后，将玻片放置晾干约 12 h，然后小心地将它们移走并按照下述的方案进行处理。晾干可以促进点样的 DNA 与载玻片的表面包被之间的结合。
- 点样完成后，取下点样针并在超声浴中清洁。按照方案 3，在乙醇中漂洗，保存和干燥。
- 避免带乳胶手套处理玻片，因为它们可能释放荧光物质。

点样后处理

点样后处理玻片的特定方法，很大程度上取决于玻片表面的包被类型（参见关于**用于点样的基底**信息栏）。然而，点样后处理过程通常包括以下几个步骤：

74

- 将 DNA 固定到载玻片表面［采取烘焙和（或）紫外交联］。
- 预洗涤载玻片以除去未结合的 DNA。
- 使 DNA 变性。
- 封闭载玻片上残留的未结合位点。
- 通过清洗除去封闭试剂。
- 储存自制的多聚 L-赖氨酸玻片，让其陈化。

在处理前后，均可以使用方案 3 中叙述的方法之一对玻片进行分析，以确保阵列的点样正确。

材料

注意：标记有〈!〉的材料的正确处理方法见附录 3。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成见附录 2。
使用前将储存液稀释到合适的浓度。

乙醇（95%和 70%）

据报道，有些批次的乙醇（100%或 200 标准强度）中含有会使微阵列试验中产生很深荧光背景的污染物质，95%的乙醇不存在这个问题。为避免这个问题，可以用 95%的乙醇或者已证明不会产生高荧光背景的无水乙醇。Warner-Graham 公司（Cockeysville，Maryland）出售的无水乙醇（USP 级乙醇，64-17-5）效果很好。

H_2O（灭菌）

$NaBH_4$（S 9125，Sigma）〈!〉

磷酸盐缓冲液（PBS），不含 Mg^{2+} 和 Ca^{2+}

SDS（0.2%）〈!〉

核酸和寡核苷酸

点样微阵列

专用装置

用来标记载玻片的玻璃刀

设定为 100～140℃的热板（步骤 3）

丁腈胶手套（无粉末）

烘箱，设定为 80℃

载玻片架（900234，Wheaton）

台式离心机，配备有玻片固定器（如 Beckman GS-6KR 或 Allegra 6R）

紫外交联仪（Stratagene 1800 或 2400 Stratalinker）

水浴锅（沸腾的）

75

下面的方法主要适合于多聚 L-赖氨酸及醛基修饰的玻片，其中很多步骤是处理不同类型的玻片时通用的。对于许多普遍使用玻片的特殊要求，见本方案后面关于**处理点样后载玻片的特别注意事项**信息栏。适合于不同商品化载玻片的处理程序，也可参考供应商的方案。

方法

标记载玻片

1. 戴上无粉末丁腈胶手套（nitrile glove），从点样平台取出玻片放到玻片固定器上。
2. 用玻璃刀在玻片的反面沿着阵列的上、下边缘各画一条线，标志出阵列的轮廓。

   在点样后的处理中，盐被除去后，阵列将变得不可见。玻片上的标记有助于正确添加靶溶液和放置盖玻片。

重新水化

3. （可选）利用温水（<37℃）浴的蒸汽，把每片载玻片有阵列的一面轻轻润湿 1～5 min，使得 DNA 重新水化。为了防止冷凝，在玻片暴露到水蒸气之前，应处在室温下。阵列上的样点应该是许多“闪光”的亮点，而不是连成一片。在温度为 100～140℃的热板上加热 1～2 s，将玻片迅速干燥。

   一部分研究人员相信这一步操作是必需的，而其他人则认为没必要。该步骤是否必须取决于点样溶液、载玻片基底和点样仪/针的性能。它的目的是希望 DNA 被均匀地分布于整个样点，减少圆环形状的出现。然而，如果操作有误，重新水化可能会导致点印的样品拖尾（smearing）。当玻片上出现许多如同第 3 章疑难问题解答指南中的圆环形状样点时，可以考虑执行这一步骤。

将 DNA 固定到基底上

将 DNA 共价固定到玻片基底上的方法通常为烘焙或者紫外交联，具体选择哪种方法取决于玻片基底的性质。因为将 DNA 共价连接到醛基修饰的玻片时，会发生氨基-醛基的缩合反应，因此点样到这些玻片表面的 DNA 不需要固定。将 DNA 固定到多聚 L-赖氨酸包被的玻片上，推荐办法是紫外交联法，而烘焙法更适合于氨基包被玻片的固定。对于某些表面，这两种办法中的任何一种均可，可以单独或是联合起来使用。

4. 将 DNA 固定到载玻片基底（关于条件、时间和设定的一般建议，见表 2-2；关

于特殊的方法，见本方案后的栏目）：

表 2-2　将 DNA 固定到载玻片基底上的推荐条件

76

玻片/玻片包被的类型	烘　焙	UV 交联
聚 L-赖氨酸	×	60 mJ
Telechem 氨基修饰玻片	1 h	（可选择）
Telechem 醛基修饰玻片	×	×
Corning Micro Technology 玻片	2～4 h	200 mJ
γ-氨丙基硅烷（CMT-GAPSⅡ）		

烘焙法固定：将玻片在预先设定为 80℃的烘箱里放置合适的时间。

烘焙时间为 1～4 h。当只选择烘焙而不联用紫外照射时，烘焙时间需要长一些，以促进交联。

通过紫外交联固定：用 60～200 mJ 的适当功率照射玻片。

适用于那些要通过交联将 DNA 共价连接到基底的载玻片包被（如多聚 L-赖氨酸包被的载玻片）。有效的交联操作非常重要，因为交联的程度将影响到信号的强弱，所以当使用新的基底时，通过实验去寻找最佳的烘焙和（或）紫外照射时间就显得非常有价值。确保紫外灯在长时间使用中不会老化。DNA 未能充分固定到玻片上是造成“彗星”的常见原因（见第 3 章）。

5. （可选的）在温度为 100～140℃的热板上加热 20～30 s（对于多聚 L-赖氨酸包被的载玻片只需 5 s），将玻片迅速干燥。

   这种处理方法通常推荐用于氨基硅烷、CMT-GAPS Ⅱ及多聚 L-赖氨酸包被的载玻片。对于氨基硅烷及 CMT-GAPS Ⅱ包被的载玻片，紫外交联后的 20～30 s 快速热处理常常可以消除“彗星”问题。快速加热后的玻片应在空气中冷却，以防破裂。最后用大量的蒸馏水清洗。

清洗和封闭载玻片

6. 室温下，在含有 0.2%SDS 的玻璃容器中，剧烈搅拌溶液来清洗玻片 1 min，然后将玻片转移到新鲜的 0.2%SDS 溶液，重复清洗步骤。
7. 室温下，在含有新鲜超纯水的容器中，清洗玻片 2 次，每次 1 min。

   这些清洗的目的是去除点样结束后残留在载玻片上的未结合的 DNA 及盐分。若没有剧烈清洗，从载玻片上脱离下的 DNA 又会在局部发生重新结合，从而导致“彗尾”（见第 3 章）。
8. 将玻片架转移到沸腾的超纯水中 2 min 使 DNA 变性，然后将其取出放到干净的玻璃容器中，在空气中冷却至室温。

   这一步骤的目的是让 DNA 变性，使其可以与靶分子之间进行杂交反应。

77

9. 在用前配制 $NaBH_4$封闭缓冲液：
   a. 在玻璃广口瓶中，将 1 g $NaBH_4$溶解于 300 ml PBS。
   b. 溶质溶解后，加入 90 ml 95%的乙醇并搅拌溶液。
   c. 将装有封闭剂的容器放置于振荡器上，放入装有玻片的玻片架并轻轻搅动，孵育 5 min。每隔 1 min 观察一次玻片表面，并通过轻轻的手工摇动将载玻片

表面的气泡除去。

醛基和多聚 L-赖氨酸包被的载玻片通常要封闭以除去表面的活性位点，否则它们可能会与标记的靶探针结合。氨基包被玻片的封闭是可选择的，取决于玻片使用过程中所遇到的背景水平。

可选择的使用琥珀酸酐封闭

据报道，含有 1，2-二氯乙烷与 *N*-甲基咪唑的琥珀酸酐溶液可以通过加快封闭过程来降低背景，且在封闭清洗过程中限制已释放的 DNA 再次结合到玻片表面。

1. **琥珀酸酐封闭溶液的配制**
   a. **在 50 ml 的圆锥形试管中，配制 15 ml 0.2 mol/L 的硼酸钠。**
   b. **将 6 g 琥珀酸酐溶解在 325～350 ml 的 1-甲基-2-吡咯烷酮中。加入试剂要迅速。**
   c. **琥珀酸酐溶解后，立即加入 15 ml 硼酸钠溶液（步骤 a），并充分混合。**
2. **迅速将封闭溶液倒入空的玻片容器中。**
3. **迅速将装有玻片的架子放入封闭溶液中，上下剧烈摇动玻片架数秒钟，确保载玻片始终都浸没在溶液中。**
4. **在有轨振荡器上，将玻片在封闭溶液中孵育 15～20 min。**

10. 用 0.2% SDS 溶液清洗玻片 3 次，每次完成后在室温下再用超纯水清洗 1 min，同时连续摇动。

    一些研究人员提议最后用 95%乙醇漂洗或者将玻片在沸水中加热数秒钟，以便在离心及储存前，更有效地干燥。

11. 将玻片放入台式离心机中，500 *g* 离心 3 min，以除去玻片表面残留的液体。将玻片储存在避光的盒子里。

    使用不含软木衬垫的塑料盒，因为软木可能会污染玻片表面。一些研究人员将多聚 L-赖氨酸包被的玻片在用于实验之前“老化”数天到一周。

78

处理点样后玻片的特殊条件

不同的点样后处理有多种选择，下面的总结适用于一些常用的玻片。

SuperAmine 方案

1. **在 80℃下，烘焙玻片 60～80 min。**

   点样阵列在烘焙后储存，在使用前立即处理。

2. **（可选择）重新水化、快速加热，在 60 mJ 下紫外交联。**

   这一系列步骤可能不是必需的，因为超过 1 h 的烘焙足以让 DNA 固定在玻片表面。然而也有报道，在烘焙后再用紫外交联可以增加 DNA 在氨基化表面的固定。由于交联剂效率有变化，要对暴露在紫外线下的条件进行优化，以适合于特定的实验室条件。

3. **用以下一系列溶液，搅动清洗 2 min。**

   注意事项：不要使用玻片固定器——它们在转移到冰冷的乙醇中时会损坏。

0.1% *N*-十二烷基肌氨酸钠

2×SSC，在室温下

95℃ H_2O 或者沸水

冰冷的乙醇

用 *N*-十二烷基肌氨酸钠来代替 SDS 是因为它更容易从玻片上洗掉。SDS 如果没完全洗掉，将会产生很强的荧光。

4. 500 *g* 离心 2 min 进行干燥。
5. 将处理过的玻片立即用于杂交。

SuperAldehyde 方案

不需要烘焙或交联。点样后的阵列晾干并储存，在使用之前迅速处理。要注意该方案中省略了封闭步骤，尽管它可以降低杂交背景，但被认为会降低阵列发出的荧光信号强度。

1. 用以下一系列溶液，搅动清洗 2 min。

   注意事项：不要使用玻片固定器——它们在转移到冰冷的乙醇中时会损坏。

   0.1% Sarkosyl

   2×SSC，在室温下

   95℃ H_2O 或者沸水

   冰冷的乙醇

   用 Sarkosyl 来代替 SDS 是因为它更容易从玻片上洗掉。SDS 如果没完全洗掉，将会产生很强的荧光。

2. 500 *g* 离心 2 min 进行干燥。
3. 将处理过的玻片立即用于杂交。

CMT-GAPS Ⅱ 方案

1. 将玻片重新水化并稍作加热。
2. 在 80℃下烘焙 2～4 h 或者紫外交联（200 mJ）
3. 将玻片重新水化并稍作加热 20～30 s。

   若没有该步骤，这些载玻片通常会产生彗尾，特别是用浓的溶液点样时。此外，过度的水化则可能导致样点扩散和样点形态上的破坏。

4. 在空气中冷却至室温（以防玻片破碎）
5. 如果背景较高，用琥珀酸酐封闭。

   这一步是可选的，处理这些载片时通常将其省略，特别是预杂交中使用了含有 BSA 的溶液时。

6. 室温下用 0.2%SDS 清洗 20 s，然后再在室温下用 H_2O 清洗 20 s。
7. 在 95℃下，用 H_2O 清洗 2 min，然后室温下在 95%的乙醇溶液中浸泡 1 min。
8. 干燥并储存玻片。

多聚 L-赖氨酸方案

1. 将载玻片重新水化和紫外交联（60 mJ）
2. 重新水化并稍作加热 5 s，然后冷却至室温。
3. 用琥珀酸酐封闭。
4. 室温下用 0.2% SDS 清洗 20 s，然后再在室温下用 H_2O 清洗 20 s。
5. 在 95℃的水中加热 1.5～3 min。

   避免水沸腾，因为它可能导致玻片包被软化并诱发剥落。

6. 室温下在 95%的乙醇溶液中浸泡 1 min。
7. 干燥并储存载玻片。

79 # 方案 3

检查印刷玻片的质量

Bi-Huei Hou, Shauna Somerville（华盛顿卡内基研究院，斯坦福，加利福尼亚，94305）；
Andrew Holloway, Maria Murphy（Peter MacCallum 癌症研究所，东墨尔本，澳大利亚）。

在与有价值的靶序列杂交之前，有多种方法可以用来评价微阵列的质量，包括用二聚体的青色素染料对阵列染色，或者使用标记的引物与阵列杂交，引物的互补序列可以位于克隆集的引物位点中或者在引物位点下游的载体序列之中。对玻片最严格的测试包括与一系列反映了不同表达谱的、截然不同的标记 RNA 样品杂交。

一些扫描仪配有两种以上的激光。在这些扫描仪中，用适合于第三种激光发射波长的通用 DNA 染料染色，第三种激光可以用来评价已杂交玻片的质量。不管使用哪种方法，这些采样过程在一次点样操作中均应该对连续的每 15～20 张玻片，采样一片进行检查，以便确保印刷玻片的质量。

微阵列也可以在点样后处理（固定）之前使用配有 633 nm 滤色镜的扫描仪通过红光反射扫描（red reflection scaning）进行评价，如后面几页中关于**阵列后处理之前对点样的 cDNA 成像**信息栏中叙述的那样。图 2-2 和图 2-3 中列举了 Yue（2001）对评价阵列不同方法的比较。

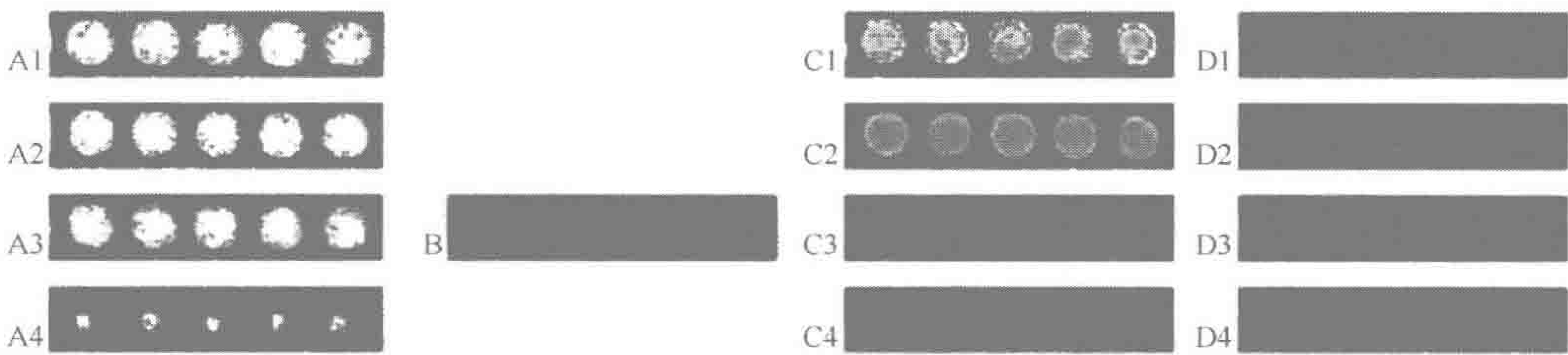


图 2-2　阵列上完好的 DNA 样点示例

浓度为 800 ng/μl、600 ng/μl、200 ng/μl、50 ng/μl 的 cDNA 样品按 5 个重复点样到玻片上，然后用红光反射扫描检查（A1～A4），ToTo-3 碘化物染色（C1～C4），或者与 Cy5 标记的 *Arabidopsis* Col-0 总 RNA 杂交（D1～D4）。图 2-2B 中，浓度为 200 ng/μl 的 cDNA 样品按照相似的方法点样，并用 PoPo-3 碘化物染色。在图 2-2B、C1～C4 中，每个样点亮度用颜色来表示。样点亮度强弱任意排列为白、红、黄、绿、蓝，其中白色最强。图 2-2D1～D4 的蓝色由于与 Cy5 标记的靶序列杂交，因此不能直接与其他图板上的蓝色作定量比较（见彩图）

80 建议在阵列的多个不同位置上，点样一组完全清楚的阳性、阴性和添加对照的样点（关于使用各种对照益处的讨论，请见第 1 章关于**用于点样微阵列的对照**信息栏）。在每次点样及杂交中，对照的表现为确定存在的问题和诊断问题的起源提供了有益的帮助。

利用荧光 DNA 染料来评价微阵列

二聚体的青色素核酸染料是一种特别灵敏的荧光探针，对核酸表现出很高的亲和

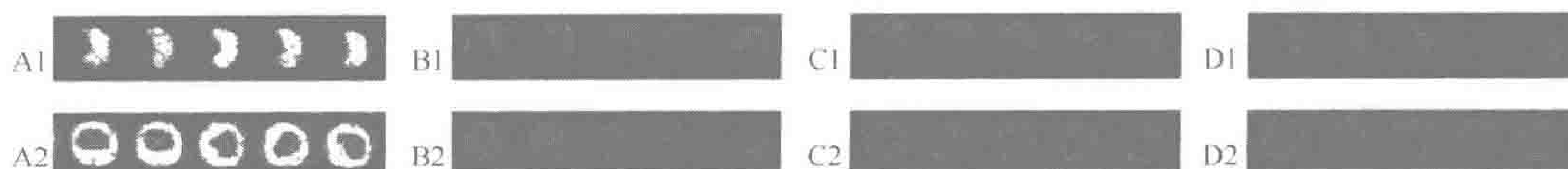


图 2-3 阵列上不完整 DNA 样点示例

浓度为 200 ng/μg 的 cDNA 样品按 5 个重复点样到玻片上，然后用红光反射扫描检查（A1、A4），PoPo-3 碘化物染色（B1、B2），ToTo-3 碘化物染色（C1、C2），或者与 Cy5 标记的 *Arabidopsis* Col-0 总 RNA 杂交（D1、D2）。不完整的样点以月牙形状（A1、B1、C1、D1）或者圆环形状（A2、B2、C2、D2）出现。在图板 B、C1、C2 中，每个样点亮度用颜色来表示。样点亮度强弱任意排列为白、红、黄、绿、蓝，其中白色是最强的。图板 D1、D2 的蓝色由于与 Cy5 标记的靶序列杂交，因此不能直接与其他图板上的蓝色作定量比较（见彩图）

力。通过使用对 DNA 亲和力很强的染料如 SYTO 61（Yue et al. 2001）、PoPo-3 或 ToTo-3 给点样的阵列染色，可以测量 DNA 沉积的均匀度以及样点的形态。这些可以用来给双链 DNA 或单链 DNA 和 RNA 染色的花菁类染料，当没有 DNA 结合时基本上不产生荧光，而当其与 DNA 结合后，荧光强度将增高 100～1000 倍。推荐使用 ToTo-3 染料，因为 PoPo-3 有时会给出很高的背景，其原因尚不清楚。当用来点样的 DNA 溶液浓度在 800 ng/μl（图 2-2，C1）到 50 ng/μl（图 2-2，C4）之间时，通过 ToTo-3 碘化物的染色，可以根据样点的强度来粗略估计所点样的 DNA 的数量。通过比较图 2-2 中 C3 与 D3，可以发现由 ToTo-3 碘化物染色得到的 DNA 样点大小与杂交得到的结果有很好的相关性。

该方案包含了一系列用于评价点样阵列质量的程序：一个基本的方法是利用荧光染料染色，一个可供选择的方法是通过与 Cy 标记的引物杂交，本方案后的信息栏中描述了将阵列与从不同细胞系或组织来源的 RNA 所制备的靶序列杂交。

阵列后处理之前对点样的 cDNA 成像

阵列上样点的大小及形状可以通过暂停点样过程并利用放大镜来直接检查，或者在点样后利用解剖显微镜观察。与单根针相关的不规则样点暗示了与针有关的问题（方案 2）。通常，较大样点含有较多的 DNA。然而，由于样点的大小取决于玻片表面的包被，因此使用一种包被的玻片上的样点与使用另一种包被的玻片上的样点是不可比的。阵列上样点的大小和形状也可以使用配有 633nm 滤色镜的扫描仪通过红光反射扫描（red reflection scaning）进行评价（如 Packard Bioscience、Scan Array Lite 4000 或 5000）。要注意这些方法既不适用于 DMSO 或其他非盐的点样缓冲液，也不适用于沉积在阵列样点上的盐晶体被洗去时的情形。

Packard Biochip（http://www.packardbiochip.com）的技术人员发明了一种被称为“反射成像”的方法，它通过扫描仪以一种简单的、非破坏性的方式来观察用含有盐的 DNA 溶液点样的玻片。该方法的原理是通过检测 DNA 溶液在玻片表面干燥后残留在每个样点中的盐晶体反射的激光来确定样点的大小和形状。反射光线在为 PMT 收集之前通过中性滤色镜的过滤。尽管含有极少量 DNA 的样点能被识别（比较图 2-2 A3 中 200 ng/μl 点样的样点和图 2-2 A4 中 50 ng/μl 点样的样点），反射图像不能用于定量估计玻片上 DNA 的点样量。根据作者的经验，从红光反射扫描得到样点大小不能与那些从杂交得到的进行直接对比（比较图 2-2 中的 A4 到 D4），因此，该方法主要用来检查每个样点的有无。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。
使用前将储存液稀释到合适的浓度。

二聚体的青色素核酸染料（Molecular Probes）

PoPo-3（P-3584，Molecular Probes）或 ToTo-3（T-3604，Molecular Probes）在本方案中均可使用。推荐使用 ToTo-3 染料，因为 PoPo-3 有时给出很高的背景，其原因尚不清楚。
提供的 PoPo-3（相对分子质量 1223）是浓度为 1 mmol/L 的二甲基甲酰胺溶液〈!〉；其吸收峰（Abs）及荧光发射峰（Em）分别为 534 nm 和 570 nm。ToTo-3 在 DMSO 中的浓度为 1 mmol/L；其吸收峰（Abs）及荧光发射峰（Em）分别为 642 nm 和 660 nm。到货时，将染料溶液储存在−20℃的干燥器中，可以使用最多 6～12 个月。在使用之前，将溶液加热到室温并彻底混匀。

10×磷酸盐缓冲液（PBS）（70013-032，GIBCO）

清洗溶液

10 mmol/L Tris（pH 7.5）
1 mmol/L EDTA
0.1% SDS〈!〉

核酸和寡核苷酸

点样 DNA 微阵列

专用设备

离心真空浓缩器（如 SpeedVac）
清洗玻片的玻璃瓶
微阵列扫描仪

一些扫描仪配有两种以上的激光。在这些扫描仪中，用适合于第三种激光发射波长的通用 DNA 染料染色，第三种激光可以用来评价已杂交玻片的质量。

塑料平板（如能放 4 块玻片的塑料盖*）（EK26161，E&K）
或
塑料盖玻片（22 mm×40 mm）（如 Hybrislip；247456，Research Products）
玻片架（900234，Wheaton）

82

方法

1. 将冷冻的 ToTo-3 染料解冻，并用 1×PBS 在室温下将染料稀释到最终浓度为

* 原书印刷错误，plastic lip，应为 plastic lid——译者注

1 μmol/L。

2. 将 50 μl 已稀释的染料滴加到塑料盖玻片或塑料平板上，然后将玻片（点有样品的一面朝下）盖到盖玻片上，确保已稀释的染料溶液立即将整个玻片覆盖。

   作为替代，向阵列（朝上）的中央加入适量体积（1.8 μm×1.8 μm 的阵列使用约 15 μl）的染料溶液，然后立即用塑料盖玻片覆盖玻片。

   注意事项：必须将稀释过的染料快速加到阵列上！

3. 室温下，在黑暗中将阵列与稀释的染料孵育 30 min。
4. 在孵育过程中，将玻片架放置在装有 350 ml 1×PBS 的玻璃瓶中。
5. 30 min 孵育结束后，用镊子将玻片放置到浸没在 PBS 中的玻片架上。在瓶里上下晃动玻片架 2 min，以除去盖玻片及未结合的染料。
6. 室温下，在真空浓缩器中（SpeedVac）以最低速离心 5 min，将玻片干燥。
7. 使用下述方法之一对 DNA 进行扫描。

   检测 ToTo-3 碘化物染色的 DNA：使用红色激光及 Cy5 滤色镜扫描阵列。与双链 DNA 结合时，ToTo-3 的吸收/发射峰为 642/660 nm。

   检测 PoPo-3 碘化物染色的 DNA：使用绿色激光及 Cy3 滤色镜扫描阵列。与双链 DNA 结合时，PoPo-3 的吸收/发射峰为 534/570 nm。

   尽管不推荐使用已被染色过的玻片，但如有必要，在染料除去以后，玻片可以用于杂交。为了除去染料，将玻片在清洗液中室温下孵育 1h。干燥后，玻片可以用于杂交。更多的信息，请查看网站 https://www.probes.com。

   ToTo-3 碘化物染色可以根据样点的强度粗略估计点样的 DNA 的量。当用来点样的 DNA 溶液浓度在 800 ng/μl（图 2-2，C1）到 50 ng/μl（图 2-2，C4）之间时，ToTo-3 是有效的。图 2-2 中 C3 与 D3 显示由 ToTo-3 碘化物染色得到的 DNA 样点大小与杂交得到的结果有很好的相关性。PoPo-3 碘化物染色基本上与 ToTo-3 碘化物染色相同。然而，ToTo-3 碘化物染色是首选的，因为背景通常较低。

替代方案：利用标记的引物杂交

通过和 Cy 染料标记的、与 PCR 产物使用的通用引物位点互补的寡核苷酸杂交，可以评价微阵列的质量。与利用荧光染料如 ToTo-3 相比，该方法提供了一种对与靶序列杂交时可利用探针的更好测量手段。任何荧光染料，如 Texas Red 均可用来标记引物。使用 Cy3 的好处是所有的商用扫描仪都配有合适的激光器来检测其荧光。使用含有 50 个碱基的寡核苷酸提高了结合的特异性。图 2-4 中给出了与 Cy 染料标记的寡核苷酸杂交的一个例子。

附加材料

Cy3 标记寡核苷酸，冻干的（5′端荧光染料修饰，购自寡核苷酸供应商）

序列必须与用来扩增克隆集的引物序列匹配或者处于引物位点下游的载体序列内。使用引物位点下游序列的好处是，寡核苷酸将只与扩增的材料发生结合，而不与载玻片上任何引物污染物结合。下面的寡核苷酸（50 mer）包含了用于 Research Genetic 人类克隆集扩增的反向引物的序列。

5′-3′Cy3GTGAG CGG ATA ACA ATT TCA CAC AGG AAA CAG CTA TGA CCA TGA TTA CGC

玻璃盖玻片（Esco 或 Menzel Glaser）

杂交及清洗材料见第 3 章，方案 20

方法

1. 在经过滤的灭菌 H_2O 中，将寡核苷酸按浓度 1 μg/μl 重悬，溶液应发出粉红色的荧光。将溶液分装成 15 μl 的等份，然后在－20℃下，储存在避光的盒子中。
2. 配置杂交需要的溶液（见第 3 章、方案 20），用 1.5 μg Cy3 标记的寡核苷酸来替换纯化的标记样品。
3. 将微阵列与标记的寡核苷酸杂交 12 h，然后按照与 cDNA 杂交使用的同样方法来清洗（见第 3 章、方案 20）。
4. 使用 Cy3 通道扫描微阵列。

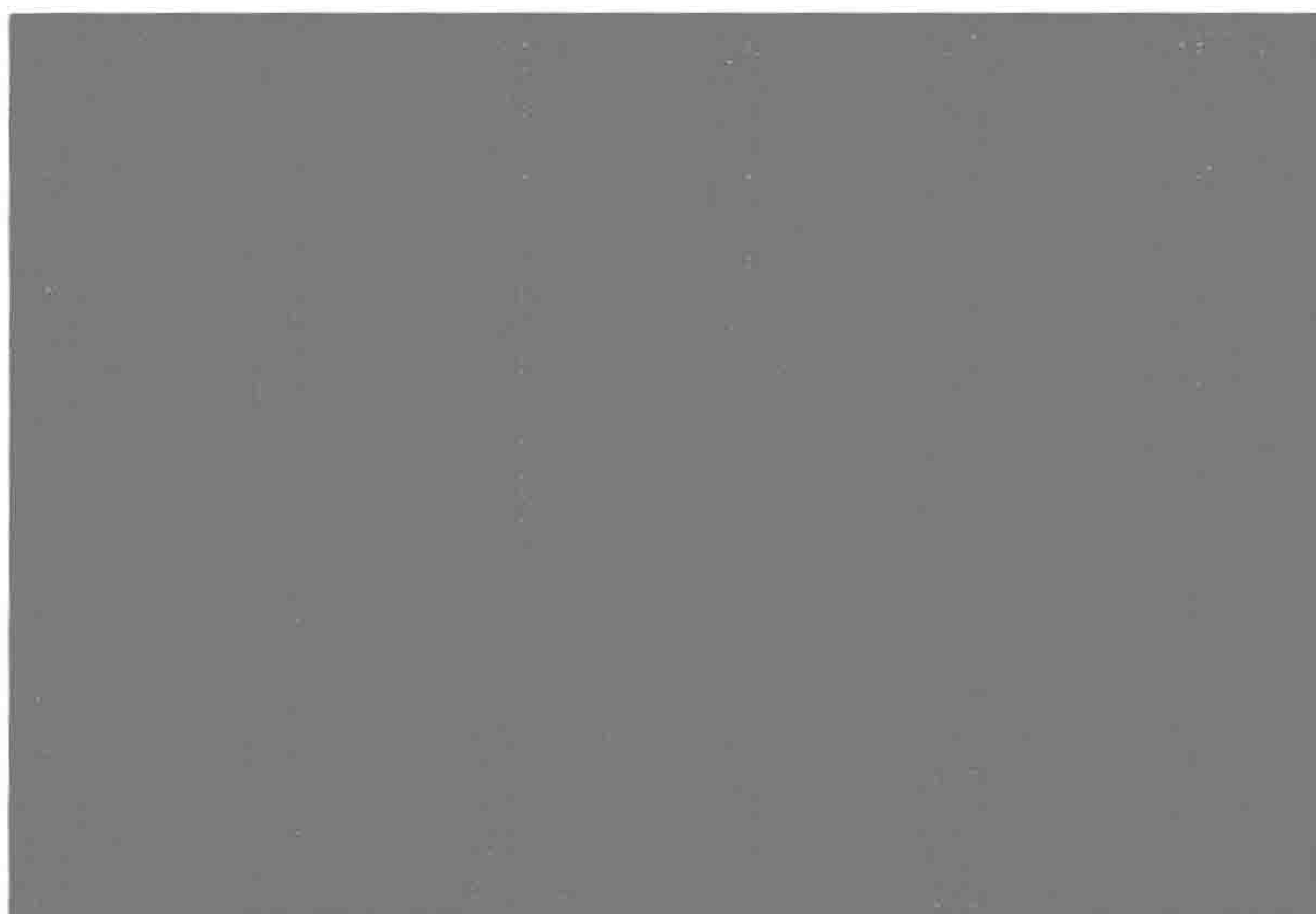


图 2-4　与 Cy3 标记引物杂交的玻片示例

几乎没有 DNA 的区域很容易看出（见彩图）

84

使用从 RNA 制备的靶序列杂交来评价玻片的质量

从表达谱有很大差别的组织或细胞类型，如乳房和淋巴细胞系（图 2-5）得到的测试 RNA 样品，可以用来和玻片杂交，以检查阵列的质量（关于标记和杂交过程，参见第 3 章，第二部分）。这些结果可以与先前使用这些测试 RNA 的实验进行比较。尽管该过程要比其他方案更加昂贵，并要求保存有一批合适的 RNA，但是它有如下几方面的优点：

- 它提供了一种为了最终目的而测量玻片质量的方法。与 DNA 具有高亲和力染料的染色不能测量 DNA 对于靶序列的可利用性。

- 如果许多靶序列被差别标记，得到的不同杂交样式有助于证实在点样过程中可能产生的错误的平板分配。与此相比，如果将玻片与荧光染料孵育或者与寡核苷酸杂交，这些错误不可能检测到。
- 检查每张玻片的极限灵敏度（threshold sensitivity），为相对以前的实验评价玻片质量提供了一种测量方法。

下面列举了建议的成对细胞系：

人类：MCF-3 和 Jurkat

小鼠：鼠的成纤维细胞和早幼粒细胞永生细胞系

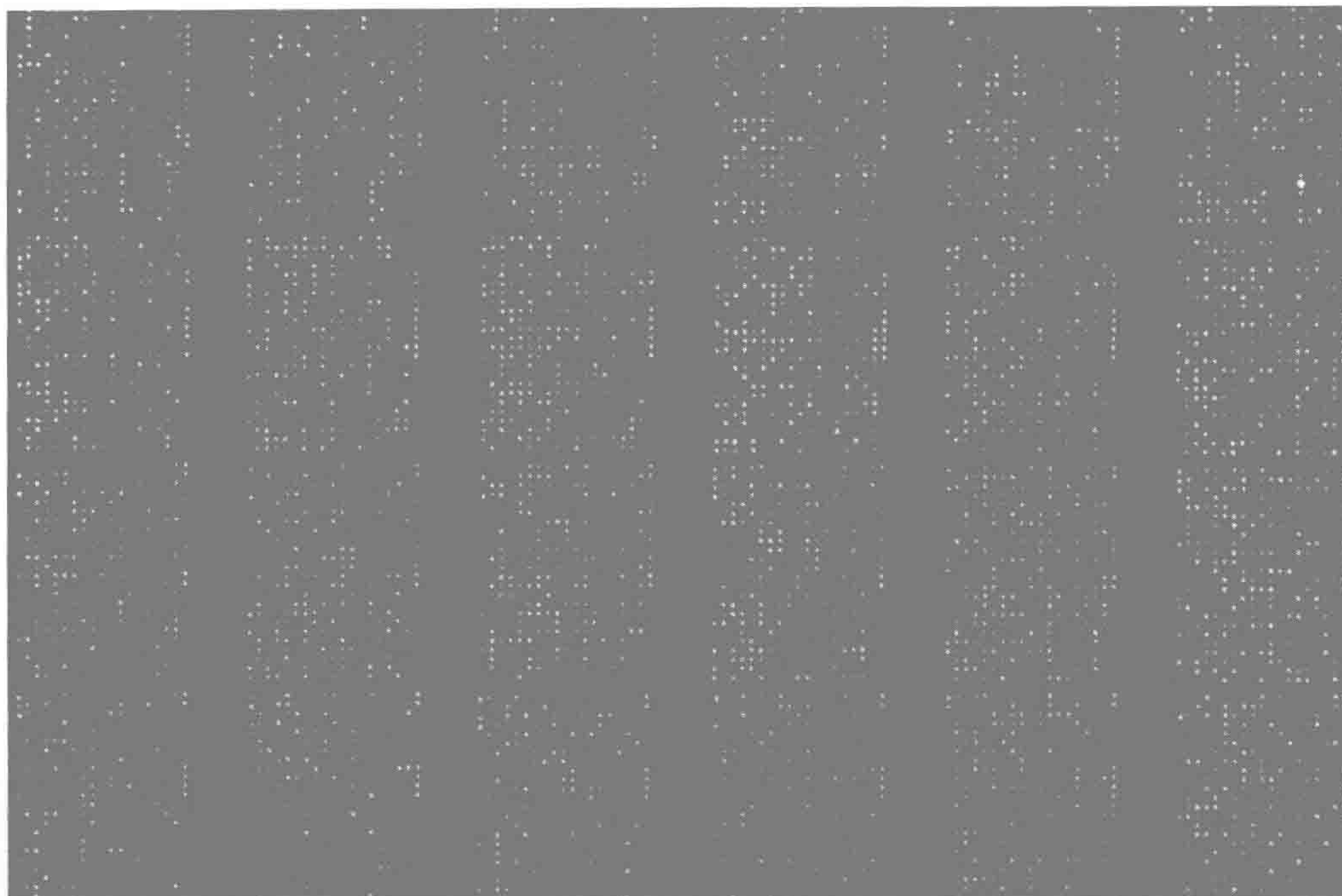


图 2-5　从 Jurkat 与 MCF-7 细胞系抽提出的 RNA 的表达谱比较

人类 10K 微阵列的伪彩色叠加图与荧光标记的、由 Jurkat（Cy3）和 MCF-7（Cy5）细胞系得到的 cDNA 杂交。得到的信号范围和杂交图谱的复杂程度可以用来比较不同批次的微阵列的质量（见彩图）

方案 4

85

鹅毛笔针的保养与清洗

Andrew Holloway，Maria Murphy（Peter MacCallum 癌症研究所，墨尔本，澳大利亚）；

Aldo Massimi，Thomas Harris[†] **and GeoffreyChilds**（阿尔伯特·爱因斯坦医学院，布朗克斯，纽约，10461）

在重新从点样板加样之前，鹅毛笔针可以用来对多张玻片进行点样。然而这种类型的点样针，针尖末端切削的高精度凹槽容易损坏和堵塞。当针的外表面被覆盖或者内部的凹槽受堵塞时，点样过程中样点的形状与大小会发生明显的改变（见第 3 章）。为了减少这个问题，建议使用高纯度的 DNA 进行点样，并要确保在点样循环中清洗站正常运转。在每次点样结束后，应将针进行超声清洗。下面的方法适合于将针取下进行清洁。在清洗针时要戴上无粉的丁腈胶手套。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

清洗溶液

下面列举了用来在超声波清洗器中清洗不同来源点样针的推荐溶液。

- 温水（Majer Precision）
- Loctite 7840 去脂剂，在水中按 1∶100 稀释（Genetix）
- 2%～5% Micro Cleaning Solution 的溶液（MCS-1，Telechem）
- 50 mmol/L KOH 溶液〈!〉；在点样前，针必须被彻底地漂洗。

Telechem 警告说去离子水可能会从针中溶出离子并损伤针，要避免使用。

乙醇（95%和 70%）

据报道，有些批次的乙醇（100%或 200 标准强度）中含有会使微阵列试验中产生很深荧光背景的污染物质，95%的乙醇不存在这个问题。为避免这个问题，可以用 95%的乙醇或者已证明不会产生高荧光背景的无水乙醇。Warner-Graham 公司（Cockeysville，Maryland）出售的无水乙醇（USP 级乙醇，64-17-5）效果很好。

如果乳胶手套接触乙醇，一些物质可能从手套溶出并污染溶液，产生绿色的荧光。

H_2O（蒸馏）

使用去离子水可能导致不锈钢针的腐蚀。对于非不锈钢材料加工的针，在潮湿时容易腐蚀，特别是当溶液里含有盐时。

5×SSC

专用设备

便携式超声清洗器（32693K21，McMaster-Carr）

该超声清洗浴只有两种设置：开或关

干燥设备（干发电吹风，或者不含油的压缩空气罐）

耳朵保护装置

在超声清洗器中固定针的浮选装置

体视显微镜

方法

1. 在超声清洗器中加入 0.5×SSC 或者推荐的清洗溶液。
2. 将针放置在合适的固定器或者浮选装置中，使其漂浮在超声清洗浴的表面（图 2-6）。

   插入针穿过固定器时要仔细，即使与软的泡沫材料接触都可能损伤针的末端。当针接触到超声清洗器的侧壁时，也能造成损伤，特别是当仪器工作时。

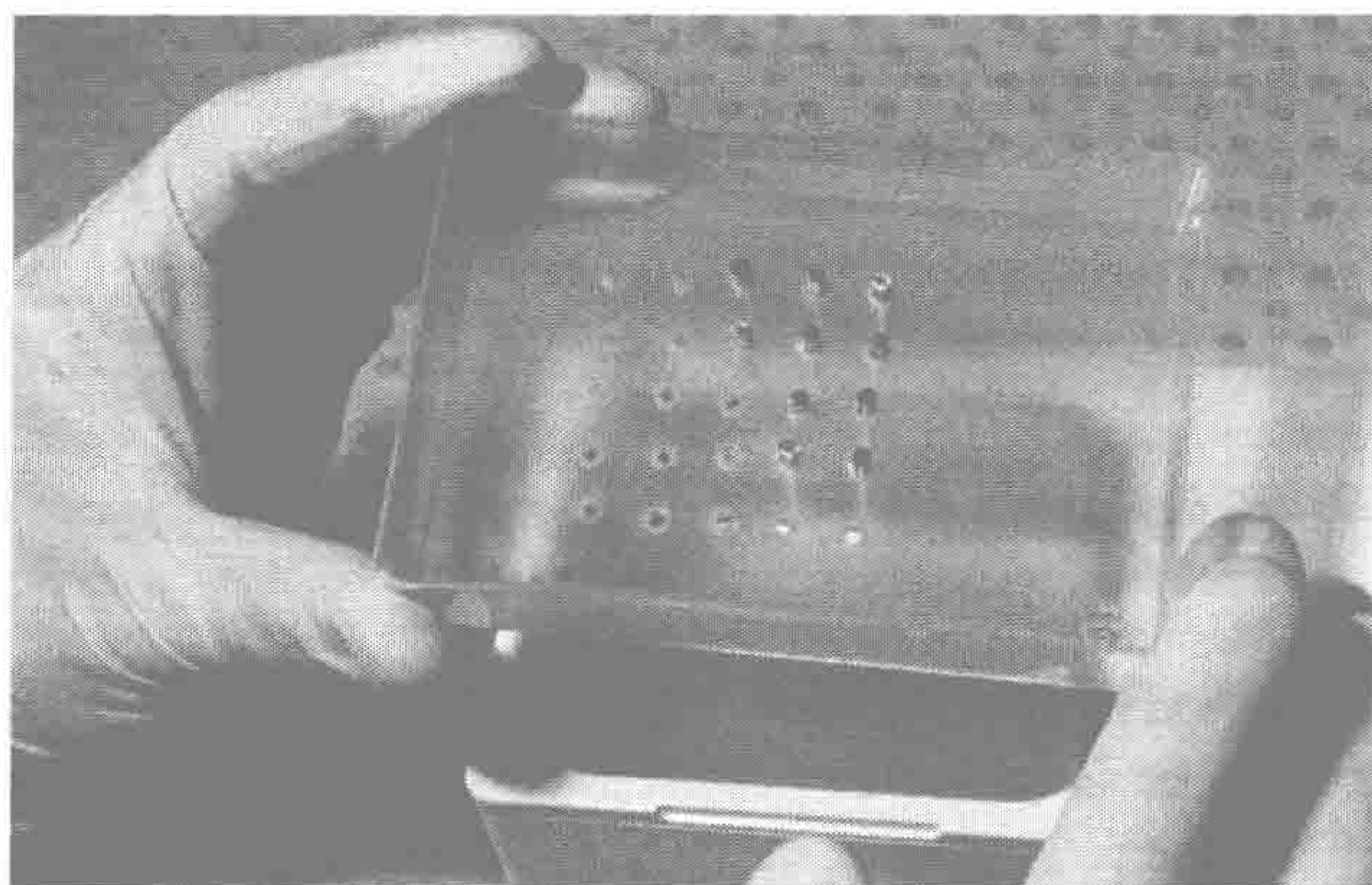


图 2-6 浮选装置

固定器用来将点样针漂浮在超声浴的表面

3. 在热的自来水中将针漂洗数分钟。
4. 在合适的清洗液中，将针超声清洗 3 min，最后用 95%乙醇漂洗。

   点样过程中的清洗程序通常是在清洗浴中循环两次，每次 3 s，接着在干燥工作站上干燥 8 s。与点样循环中针的清洗程序不同，平时清洗针所用的最适超声时间及功率设置要根据经验来确定。
5. 用干发电吹风机或其他干燥设备使针干燥。
6. 在体视显微镜下检查点样针，寻找分叉部分之间的连接物。

   在解剖显微镜下检查针时，轻轻将针的末端与一小滴 3×SSC 溶液接触。如果针工作正常，溶液可以通过毛细作用自由地进入针的凹槽中。如果针发生了严重的阻塞，按照下面关于清洁堵塞的针的信息栏中所描述的，用不锈钢薄片来清洁针尖。或者也可以将针的凹槽充满溶液，然后在显微镜下观察。如果用一定的角度来夹持点样针并且针尖向下，随着溶液干燥（如果用热的灯来加热，蒸发速度将加快），反向毛细作用行为将很明显。随着液体的干燥，从上到下空穴变空。如果在针完全变干之前，液面就从针尖移走，这样的针在点样中表现会比较差。这一现象常常发生在具有不均匀缝隙的针中。
7. 更换点样仪印刷头上的点样针。当把针的末端放入夹具时，要特别小心。针始终要在干燥条件下保存。

清洁堵塞的针

对于遭受严重污染的针，小心地用薄钢片将污染物去除或使之变松。Majer Precision Engineering (www.majerprecision.com) 建议使用以下清洗程序：

如同主要方案中所述，在温水中用超声处理清洗，并用 0.025 mm 的不锈钢薄片材料清洁，这两种方法结合起来，可以清洗堵塞的针。薄片材料可以从 MSC Industrial Supply Co. 购买 [产品号：02467751；www.mscdirect.com 或者 (800) 645-7270]

由于薄片材料很薄，所以容易剪成馅饼形状。小心将薄片材料滑进尖端凹槽的顶部，慢慢的将其轻轻移向针的末端。如果薄片楔进或卡在凹槽里，不要用力迫使它移动到末端。离尖端越远，残留物质的聚积使得凹槽变得越宽。如果聚积很严重，从凹槽的侧面仔细地用薄片材料松动残留物质，直到薄片能够在全部凹槽移动，然后用超声浴将变松的残留物去除。

88 信息栏：微阵列接触印刷使用的点样仪与点样针

接触式点样设备是到目前为止微阵列制备中最常用的系统。安放于点样头中规则排列的点样针从加样板中收集少量的探针，然后将纳升级的液体转移到显微镜载玻片上离散的位置。一些商品化点样仪的对比信息见表 2-1，图 2-7 中比较了各种点样仪器在定制的微阵列设施中的应用情况。

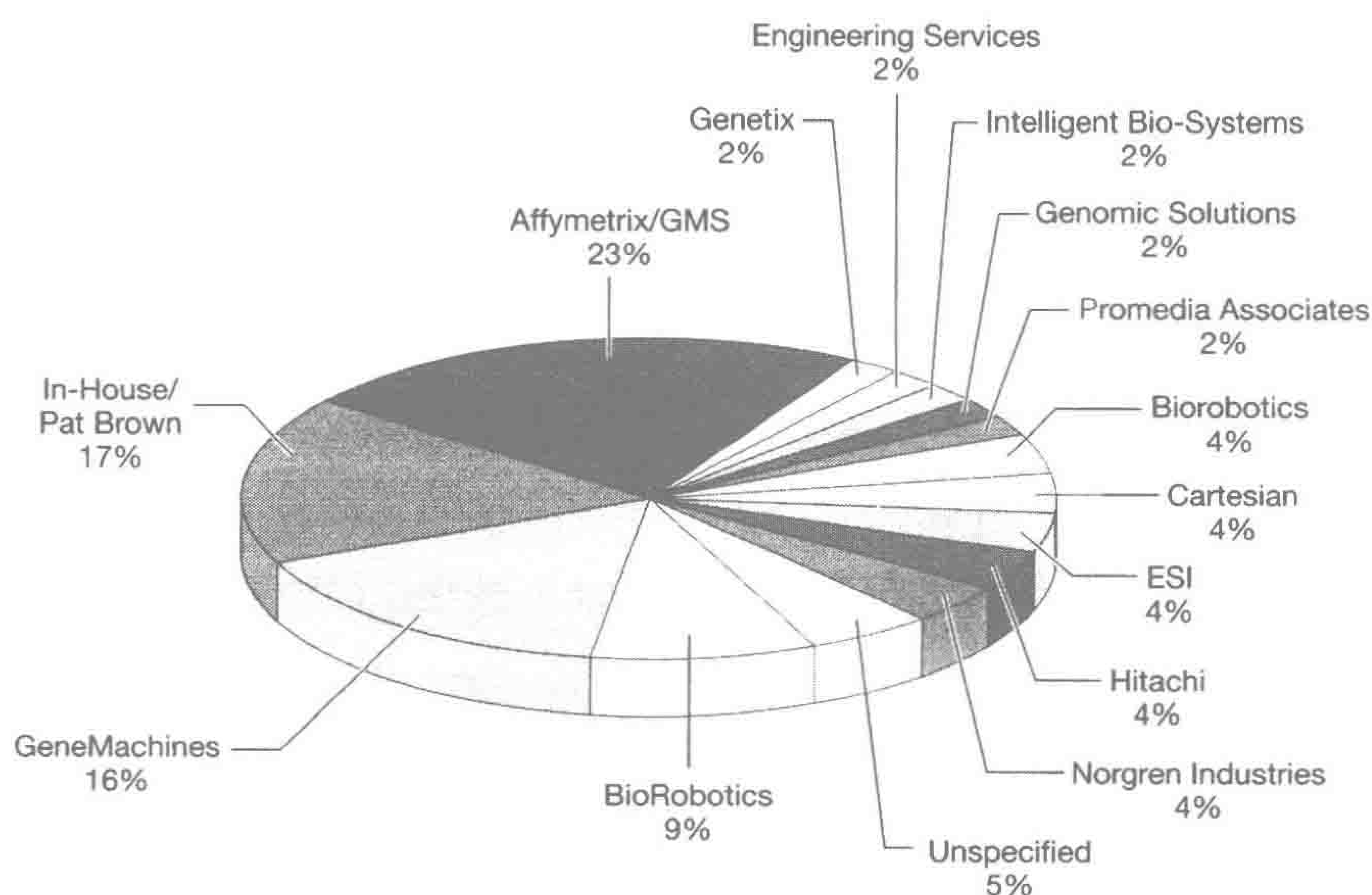


图 2-7 点样仪使用情况的比较

图表中显示了不同品牌的玻片点样仪在约 50 家用户的微阵列设施中的相对使用情况。[根据 2000/2001ABRF Microarray Research Group Study (http://abrf.org/ABRF/Research) 重新绘制]

点样仪的特性

下面叙述了基于点样针的自动点印装置的组成部分。点样仪和它的各种特点的示例见图 2-8。

- 稳定和坚固的点样平台为点样机提供了坚固的支撑和有效的减振性能，且表面的平整度可达 100 μm 或者更好。
- 一个三轴的机械自动臂提供了非常精确的定位，在台面的任意位置其重复误差均低于 10 μm。用于定位的策略包括将印刷头依次在一系列放置或固定在平台上的玻片上方移动，或者在印刷头下方移动装载了玻片的移动式平台。尽管这两种方案均是有效的，但后者具有简易性和成本优势。例如，一个配有能够同时访问 300 片玻片的上位架构台（overhead gantry）的机械臂，与配有简单的线性伺服手臂相比，成本更高、速度较慢，而且结构更加复杂。

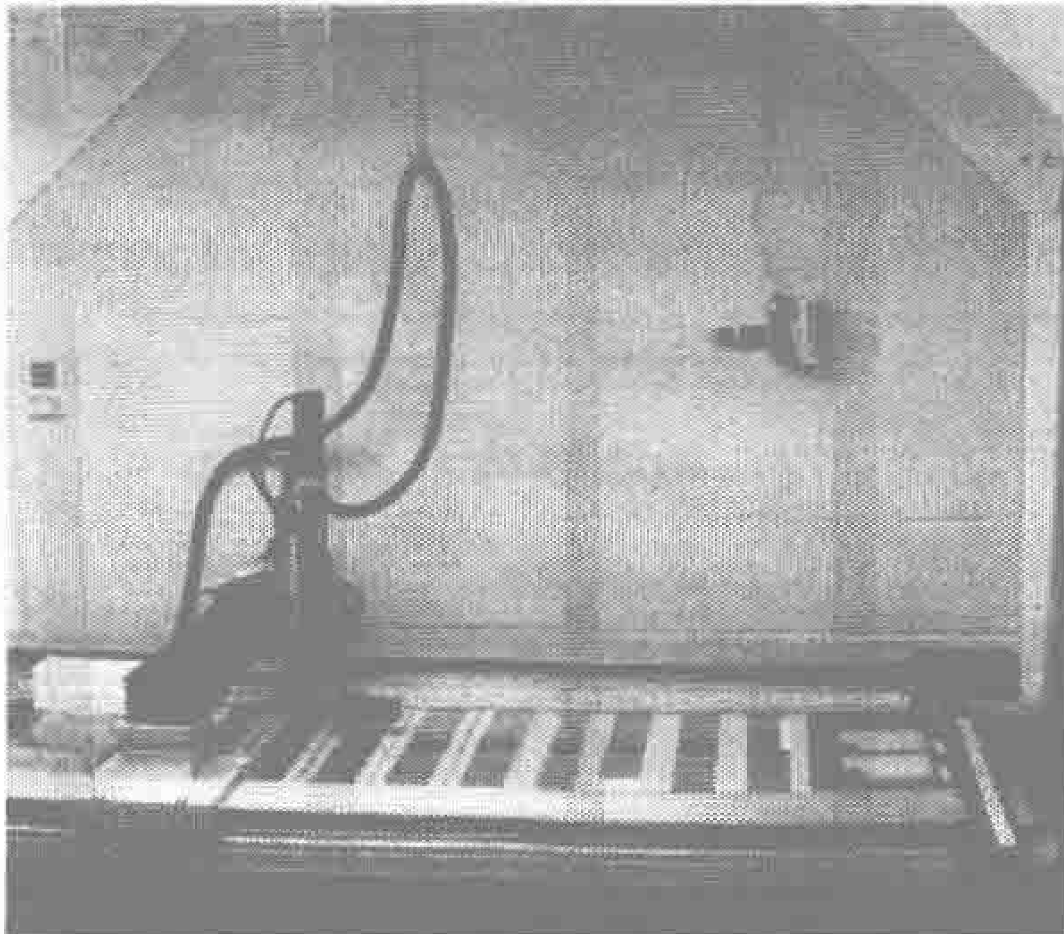


图 2-8　点样仪

图示了典型的基于点样针的玻片点样仪的不同组成部分

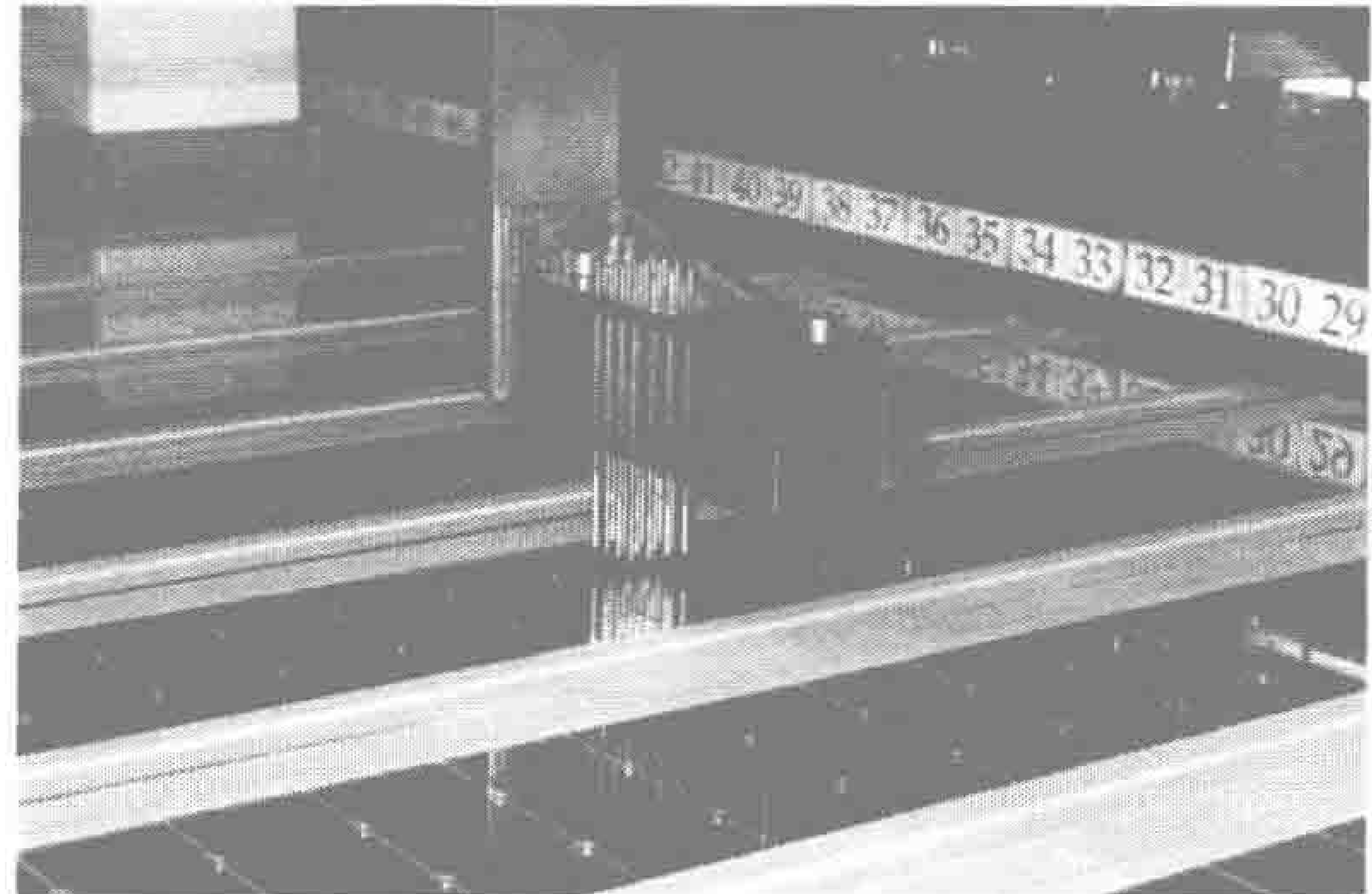


图 2-9　载玻片

在桌面上，玻片的布局和排列

- **伺服控制器**　其精密的电子元件要能够精确地控制放大器和驱动机械装置的马达。伺服控制器由高速数字电路板组成，可以放在计算机或放大器中。不同的传感器可以与控制器接口来监控机械的运转及人身安全互锁装置。
- **计算机控制**　系统通常由运行定制应用程序的个人计算机来控制。软件允许操作者通过指定组件的位置定义操作参数及计划事件的次序来设定点样过程。计算机将高级命令传输到伺服控制器，通过监视运行状况与性能来担当管理者的角色。
- **多孔板的加载**　需要有一种能精确地将一块或者更多的 96 或 384 孔板安放到平板上，并且精确度约为 100 μm 的方法。急需有一种灵活的结构，能通过侧向平移来矫正对准错误。可以将来自“旅馆结构（hotel structure）”叠放的多孔板供给点样仪的自动机械臂，可以用一种完全自动的模式来处理大量的多孔板。
- **玻片的加载**　必须将显微镜载玻片固定在台面上，以便在超过数天的时间内精确打印数千个的样点。多种方法可以使用，包括用普通的实验胶带固定，有弹簧约束的机械狭槽及带有真空排管的平台和磁铁（图 2-9）。

90

• **清洗和干燥工作站**（图 2-10）　每次点样循环后，清洁针就可以使用同一组针去点印不同的克隆，而没有明显的交叉污染。清洁至少包括用水清洗与随后的真空干燥。一些点样仪配备了超声清洗浴，从而提供了更严格的清洁。更多有关鹅毛笔针的详细清洁过程，请见方案 4。图 2-11 给出了一个污染测试结果。

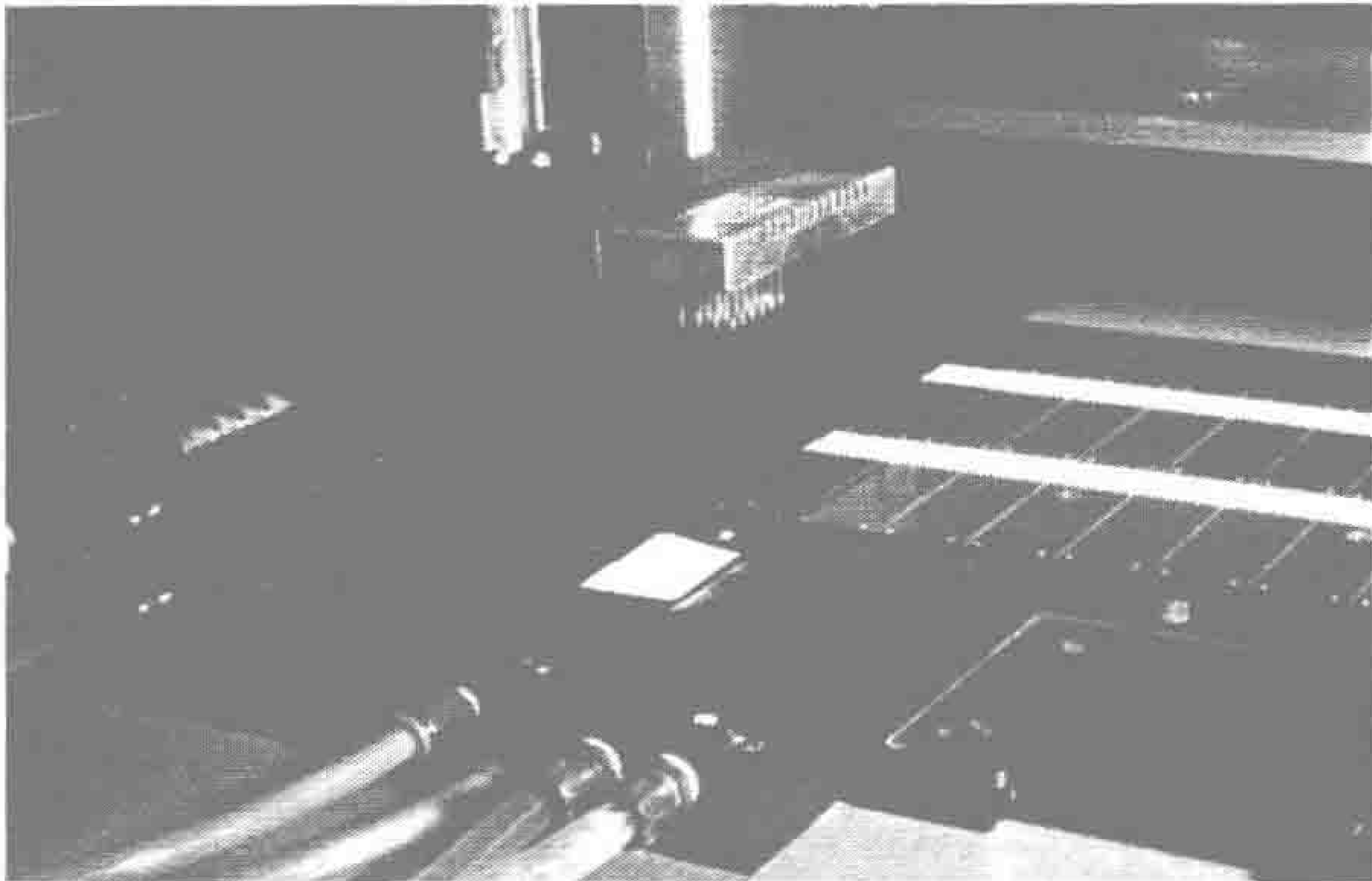


图 2-10　清洗和干燥工作站

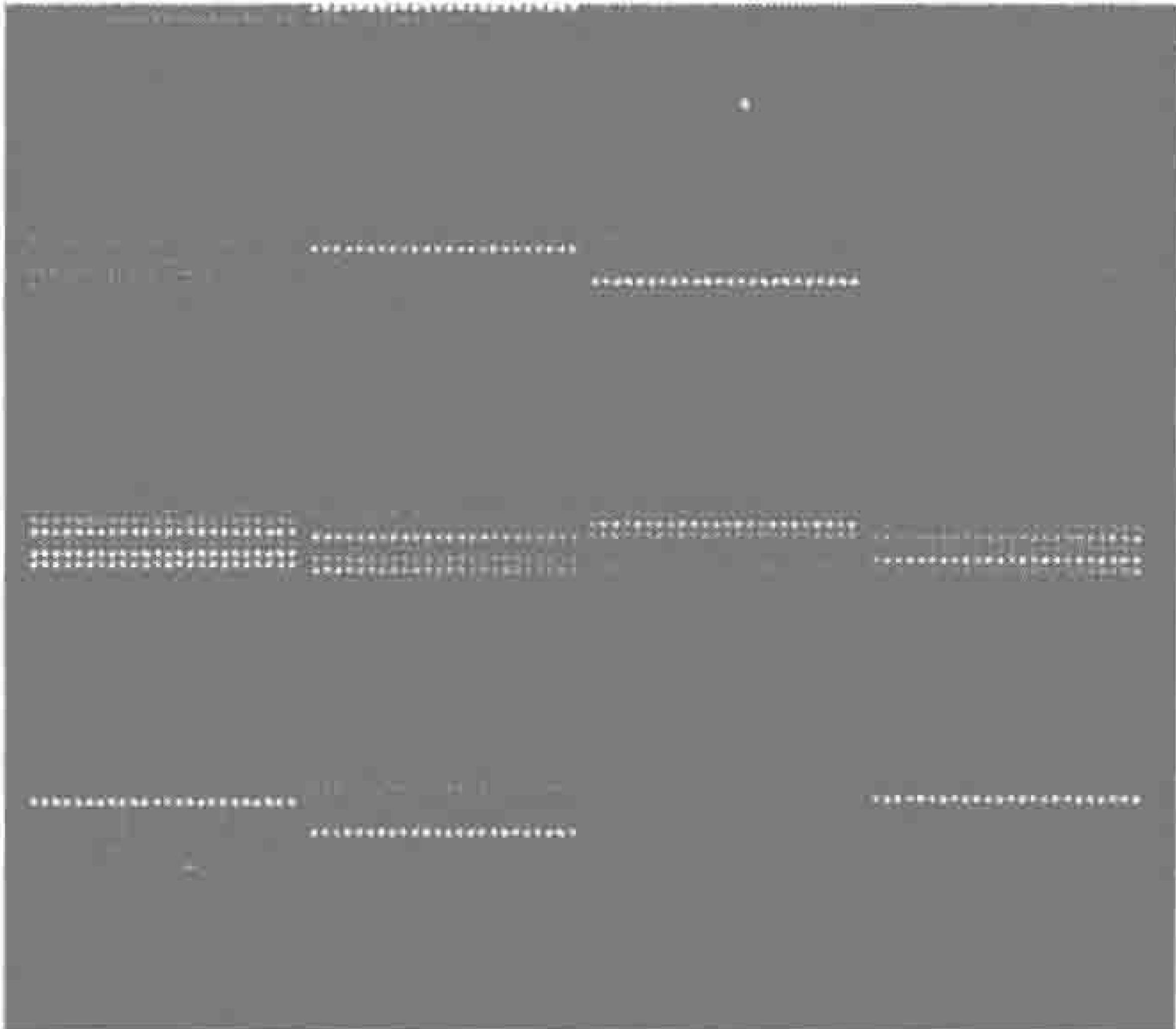


图 2-11　污染测试

选择一系列的基因来反映在一个较大阵列表面的杂交（包括颜色和强度）变化。这些克隆随后被再次点印成阵列，每根针在清洗前点印 24 个样点。接下来每根针将要点印的克隆被挑选为与先前所点印的克隆在杂交后显示不同的颜色或强度。对于 16 根针，该过程重复 5 次。最后，用针点印一行 24 个点样缓冲液的样点。对结果的评价显示，没有来自先前点样的 cDNA 克隆的残留物质，并且点印的一行缓冲液与局部背景没有明显差异。通过与 Cy3 标记的 Jurkat RNA 及 Cy5 标记的 MCF-7 RNA 之间的竞争杂交来评价玻片。实验结果与此类似，表明了常用的针清洁程序（和在方案4中描述的）可以有效地去除针上未点样的 cDNA 溶液（见彩图）

91

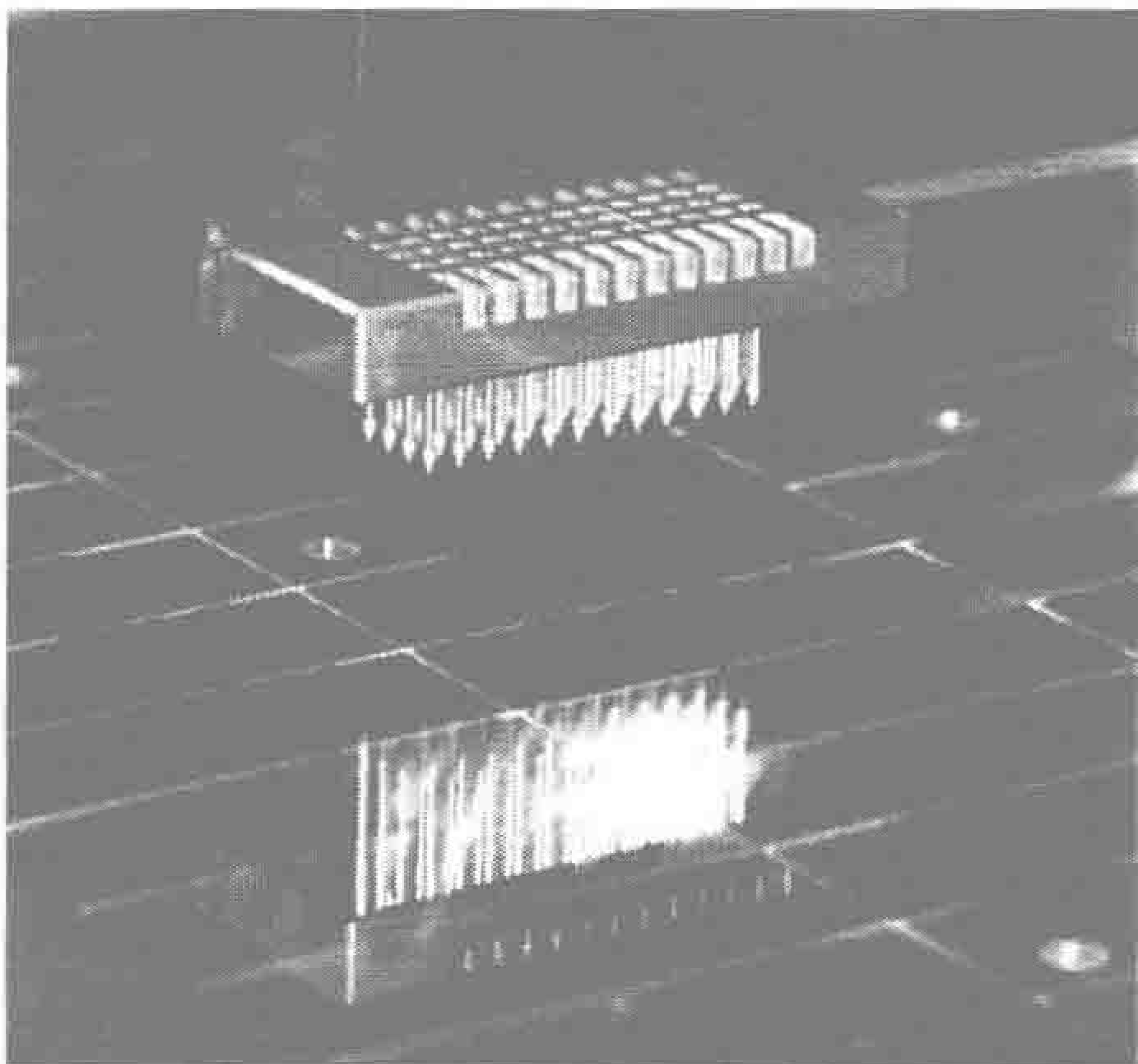


图 2-12　印刷头

• **环境控制**　周围空气中的微粒很容易污染点样的玻片，并且如果这些微粒落在点样区域，可能会导致杂交与扫描过程出现问题。因此，点样过程要在一个有层流的、HEPA 过滤空气的房间或密闭环境中进行。应该持续监视和控制空气湿度。如果湿度太低，则可能在点样循环结束之前，溶液就在针的尖端干燥。低的湿度也增加了加样板中溶液的蒸发速度，影响了 DNA 的吸取，并增加了 DNA 的浓度。每次循环结束后干燥残留在加样板中的 DNA 溶液（这样使得干燥的量不会被累积）或在下次点样前重配样品可以部分抵消这一问题。

过度潮湿（>65%）会导致样点扩散成过大的尺寸，并呈现不规则的形状（Hegde et al. 2000）。在高密度微阵列的点样中，高湿度会导致样点大小多变，造成相邻的点发生交叉污染。点样的理想湿度随着点样溶液的组成而变化（见

关于**用于点样的基底**信息栏）。然而，对于大多数常用的点样溶液，温度约20℃、相对湿度50%～60%为较好的点样条件。点样玻片上中等的蒸发速率使得DNA有时间在整个样点上均匀扩散，减少了“圆环形状”样点的形成，并改善了样点形态的均一性。一些点样仪的自动机械臂在一个密闭的点样空间中运动时会产生较多的热量，因此点样过程中的温度监测非常重要，而不仅仅是在刚刚开始时。

- **印刷头**　微阵列机械装置最终的效应器是印刷头，图2-12中给出了一个例子。该装配可以夹持不同数量的点样针进行精确排列，以适合吸取点样材料的板孔。图中显示的是48根针，其间距为4.5 mm（用于384孔滴定板），按适合于显微镜载玻片的4×12方式排列。在保持精确的侧向定位的同时，针可以靠重力或装有弹簧来保持垂直运动。限制针在点样头上的旋转是有益的。如果针发生了弯曲，它们的旋转将导致样点间隔距离的错误（见第3章中的例子）。点样过程所需的时间与印刷头上针的数目成反比。然而在一定程度上，增加针的数目与点样无直接关系，因为随着针数目的增加，遇到针的损伤和堵塞，从而危及整个阵列质量的概率也相应增加。此外，针的数目越多，诸如印刷头掉下或者与加样板对准错误等灾难带来的代价也越高。因此应该谨慎地按比例提高针的数量。针的数量必须是加样板板孔数量的整数。例如，一个4×5的针阵列不能用于384孔板，但排列为4×4或4×8的就可以使用。下面给出了有关针设计的更多详细信息。 92

影响点印质量的点样仪工作参数 93

商品化点样仪都使用工厂设定参数编程。然而，大多数仪器的设定可做调整以获得最佳的点样性能。

- 施加在显微镜载玻片表面的力和速度应该足够慢以使得针尖的变形最小，但要足够给予毛细管内的液体向下的加速度。如果接触速度太快，针尖和（或）玻片基底的表面有可能就会遭到破坏（见第3章）。尽管一定量的超行程（over-travel）是必需的，但接触到玻片时，弹簧被压缩的距离超行程应该最小化，以确保不管玻片厚度的变化和平台的起伏，每根针和每张玻片都能发生接触。选用长度一致的针尖，可以限制所需的超行程的量。
- 样点的形态取决于针中携带材料的量和针与玻片的接触时间（停留时间）。停留时间越长，点样留下的材料就越多。当针从加样板板孔里吸取点样材料出现困难时，选择在加样板孔里停留较长的时间，以增加吸入点样针中液体的量。一些点样仪允许独立调节针在玻片表面的停留时间。这就可以粗略控制递送材料的量，也就是样点的大小。
- 环境控制（温度和湿度）在上面已讨论过。

　　其他影响样点形态的因素，包括点样材料的组成和玻片表面的特性在方案2和关于**用于点样的基底**信息栏中讨论。然而，正如下面所讨论的，针的设计可能是影响点样速度、样点均一性、样点大小和样点形态最为关键的因素。

点样针的设计

- **实心针**　最简单的点样设备携带了成排的实心不锈钢针，它结实耐用，高度均一，相对容易制造（图 2-13）。钝头使得它可以在不同的基底上进行点样，包括玻璃和半多孔的基底如尼龙包被的玻片。在点样循环间，实心针更容易清洁。然而，简单的实心针一个主要的缺点是每访问加样板一次，只能进行一次点样。使用简单针的点样仪，由于缓慢的点样循环和印刷头在加样板与玻片之间连续运动会产生热量，加样板的蒸发就成为一个问题。与其他类型的针相比（100～150 μm），实心针所产生的样点直径通常较大（200～300 μm）。BioRobotics（www.Biorobotics.com）提供在他们的点样仪器和其他仪器中使用的实心针。

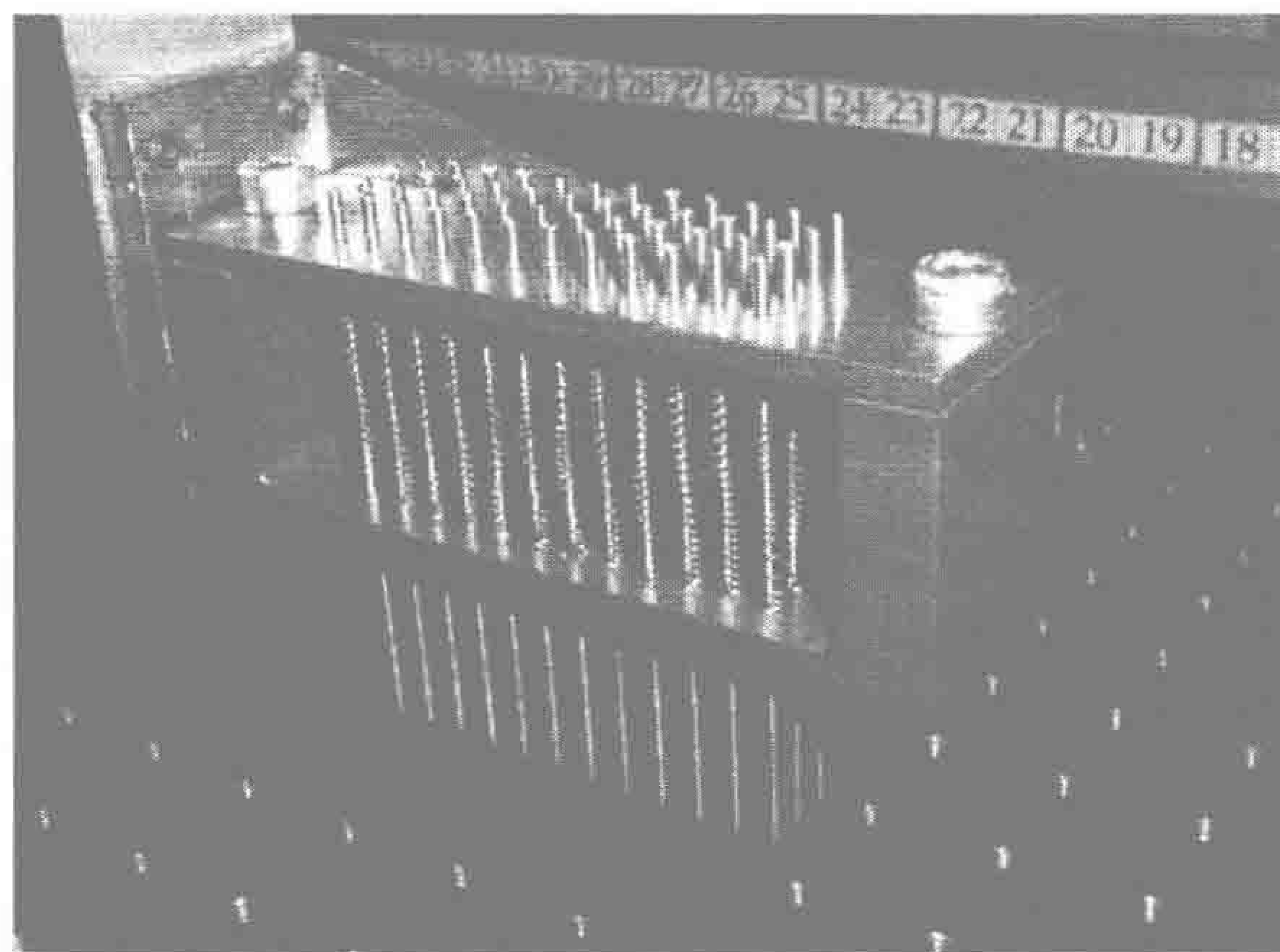


图 2-13　实心不锈钢针（详见正文）

- **鹅毛笔针**　这类针通过毛细作用填充，发明的意图在于每次从加样板吸满液体之后，可以对玻片进行连续点样。这种针带有宽度为 25～50 μm 的沟槽或狭缝，并在针尖上切削了一个蓄液池，可以让纳升量级的探针溶液吸入针尖。当装有点样液的针接触到阵列表面时，所产生的毛细作用使得探针溶液形成样点沉积。点样针加满样后，在最初的几次接触中，DNA 溶液倾向于冲出针尖。因此在开始点样前，针尖要在空白的载玻片上吸墨（blotted）。鹅毛笔针点样的大小为 90～250 μm，从而使它们适合于高密度微阵列的制备。每次访问探针溶液后可以点印的样点数量取决于针上切削的蓄液池的大小、点样时的湿度和溶解探针所使用的溶液。随着鹅毛笔针在设计上的改进，有用点样玻片的数量通常是受到点样平台上玻片容量的限制，而非最后一片玻片点样结束时可用的点样材料的量。一些针的设计使得每次装样后可以印刷大于 200 张玻片。

　　通常，鹅毛笔针不用于对半多孔基底进行点样，因为它们与滤膜接触时，蓄液池里的溶液由于毛细作用会被带走，而且锐利的针尖还会损伤脆弱的表面。使用鹅毛笔针的缺点之一就是它对损伤与堵塞非常敏感。微粒在毛细狭缝里滞留会成为非常严重的问题，导致大而不规则的样点、污染、DNA 的不均匀流动

和堵塞。此外，由于点样溶液保持在凹进去的狭缝里，不同样品间的仔细清洁就变得非常重要，以防止交叉污染或不同加样板的板孔之间的残留。针的污染程度非常依赖于点样过程中清洗循环是否充分以及点样溶液的组成。特别是残余的蛋白质（酶）或者未能重悬的 DNA 微粒会迅速堵塞鹅毛笔针。清洁通过清洗、真空干燥等循环完成，一些点样仪中还使用超声清洁。方案 4 中提供了彻底清洁鹅毛笔针的详细过程。BioRobotics（www. BioRobotic. com)、Telechem（www. ArrayIt. com）和 Majer Precision Engineer（www. majerprecision. com）生产的针在一系列接触式点样仪中被广泛使用。有关部分鹅毛笔针设计的全面的综述可以从 Brown 实验室的网站（http://cmgm. stanford. edu/pbrown/mguide/tips. html）中找到。图 2-14 显示了来自 BioRobotics 的 MicroSpot quill pin（split pin，分叉针）点样针。

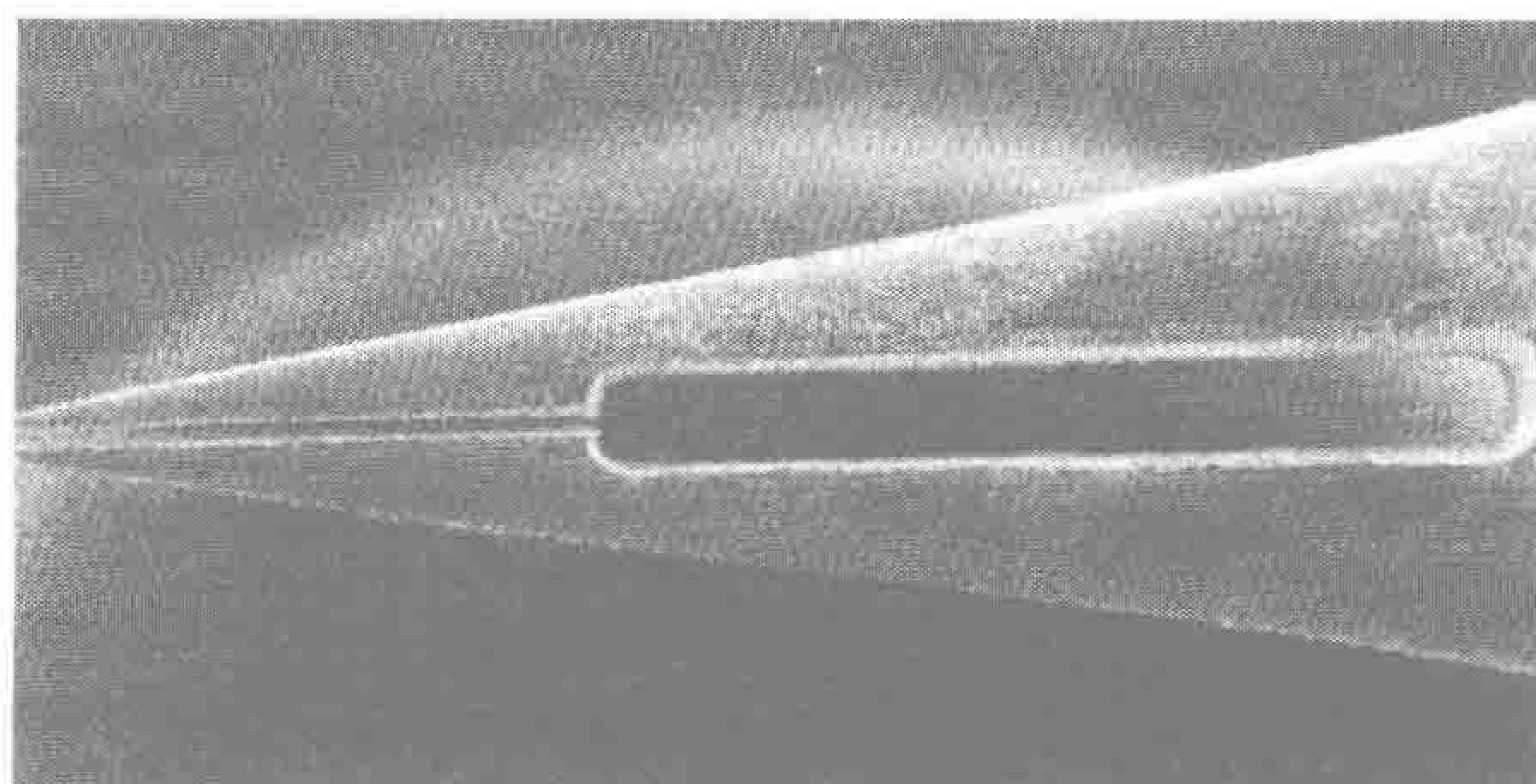


图 2-14 鹅毛笔针

图中显示的是 MicroSpot split pin［来自 J. Pearson et al.（2001）SPIE 4407，281］（详见正文）

95

- **环和针系统** Affymetrix 的 cDNA 微阵列点样仪（www. Affymetrix. com）使用环和针的组合来进行点样（图 2-15）。环作为从加样板采集的 0.5～1.0 μl 探针溶液的蓄液池。

  实心针通过环所支持的探针溶液的弯月面的中心，将相同量的溶液递送到阵列表面。弯月面的瓦解将使点样过程中止。尽管环和针的系统需要消耗大量的探针溶液，但与鹅毛笔针相比有许多优势。它具有实心针的优势，包括在尼龙膜上进行点样。尽管从理论上讲每次加样可以点印数百个样点，点样仪（译者注：应为 arrayer，原书误作 array）的载玻片容量可能使得这个优势无法发挥。环和针的组合不需要吸墨水操作，而且受点样溶液里的微粒或者不同点样溶液黏度差异的影响较少。因此，针和环系统的点样仪可使用 PCR 扩增后未经纯化的溶液进行点样（见第 4 章，方案 1）。由于清洗支持点样溶液的大孔环相对容易，因此残留几乎不再是问题。可用的最小的针为 125 μm，产生的样点尺寸为 150（±10）μm，并沉积约 50 pl 的材料，这样每平方厘米的样点密度从 1000 到 2500 个样点（或者每张玻片 10 000～25 000 个点）。300 μm 和 500 μm 的针可以用来递送较多量的材料，但该系统主要用作蛋白质点样。

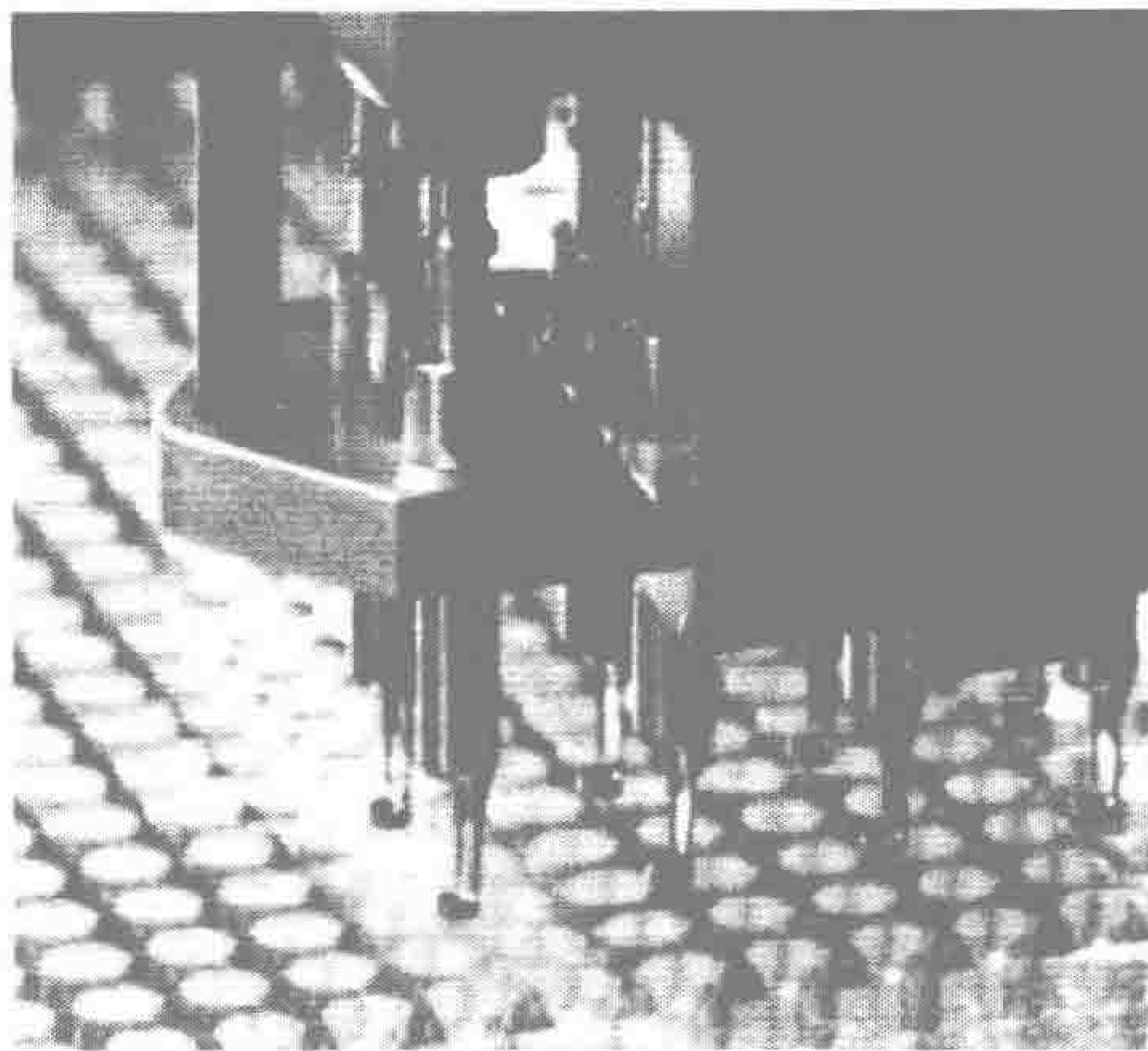


图 2-15 针和环
图中显示的是 Affymetrix 使用的针和环系统（详见正文）

96 点样针材料

点样针要使用坚硬而且耐用的金属，这样在微阵列点样的高盐环境中才不会被氧化。针通常由硬质的不锈钢制造。当浸入到滴定板孔中时，理想情况下鹅毛笔针应该只在毛细狭缝里吸取溶液，而不会黏附于外表面。表面的粗糙增加了黏附于针外表面的DNA 溶液量，影响样点大小的变化，并且由于样点增大，也浪费了点样材料。此外，表面的粗糙也促使了针上块状 DNA 斑点的形成。因此，针的外表面应该尽可能的光滑，对 DNA 溶液无亲和力，并且不应该为疏水的或带电的。如果针的尖端太锐利，由于向载玻片的毛细传递较差，开始产生的样点都会很小。锐利的针尖也更容易损坏和损坏玻片基底。在被损坏的基底上，DNA 不能与这些刺孔结合，并且当玻片杂交时，样点将呈现“笑脸”特征（见第 3 章）。因为太尖锐或者两个分叉稍微不平，新的针可能性能较差。最好使用新针进行测试点样，以便通过磨损使其尖端变得稍钝和更加均一，使毛细行为得到改善。

97 信息栏：用于点样的基底

由于玻璃载玻片具有容易获得、内在荧光较低、透明度高、热性能良好、刚性优良和易于表面化学修饰的特点，因而是受欢迎的固定探针的固相支持物。由于玻璃的无孔特性，标记的靶标可以不受内部扩散的限制直接与固定的探针接近。

常用的显微镜载玻片和特殊加工的玻璃载玻片（如那些可以从 www.arrayit.com 获得的）都可以用作支持 DNA 微阵列。多种类型的包被提供了固定点样 DNA 的表面。这些包被牢固地结合在玻片表面，并与点样在表面的 DNA 紧紧地结合。结合可以是非共价的（如疏水作用或静电吸引）或者共价的，后者可能在 DNA 分子全长产生或在末端固定，例如，将 5′端氨基修饰的寡核苷酸与游离醛基的包被结合。尽管这种定向固

定可能为靶标提供了更好的接近途径，但在微阵列实验中寡核苷酸与经酸简单清洗的玻璃载玻片也能很好地结合（Call et al. 2001）。这种情况下结合的机制还不清楚，可能是通过稳定的氢键或静电相互作用发生。不管采取何种方式结合，重要的是结合到包被的 DNA 要保留很强的与靶序列杂交的能力。

在本书写作时，关于最合适的载玻片表面，在研究人员中间还没有达成一致的意见。聚 L-赖氨酸是研究单位内部用于玻片包被最常用的材料。它不但价格低廉，而且易于得到，与商品化生产的玻片相比，后者有时订购要花上数个星期。无论玻片的来源如何，重要的是要检查它们是否具有全部规则的尺寸，这样当放置到点样平台上时，才能与印刷头保持对准。即便是为微阵列使用而特别制造的载玻片也不是总能具有规则的尺寸（关于例子，见方案 2 中的图 2-1）。

一些批次的载玻片——购买的和自制的聚 L-赖氨酸包被的玻片——偶尔会有高的背景，以荧光斑点或者荧光"薄雾"区域的形式出现。因此，建议在点样前对一定代表性数量的玻片进行例行扫描，以确保没有荧光区域干扰阵列的分析。

下面介绍了一些种类的载玻片，包括商品化、性能稳定而且品质高的载玻片。

聚 L-赖氨酸

聚 L-赖氨酸包被由于容易制备、方便得到和性能良好，而被用于第一个 DNA 微阵列的制备（Schena et al. 1996），并且仍然为大家所喜爱。DNA 结合到聚 L-赖氨酸包被的过程非常复杂，但本质上为了电荷的互相作用，并且通过烘焙或 UV 照射可以转化为共价结合。一个好的聚 L-赖氨酸包被的玻片，正确地封闭通常产生较低的背景信号和良好的信号强度。然而，聚 L-赖氨酸不是耐久的，会被点样针刺破，在玻片制备的变性过程中如果被沸腾处理会发生剥落，并且在杂交中特别是在高温下容易损伤。包被脱落暴露了未封闭的表面，会导致产生背景痕迹。

在使用前，将聚 L-赖氨酸包被的玻片储存数周，该过程称为老化，它可以预防以上的一些问题。不建议使用储存期超过四个月的聚 L-赖氨酸包被玻片。有关聚 L-赖氨酸包被玻片的制备，请见方案 1。

供应商：

Menzel-Gläser：http://www.menzel.de/menzel/index/html

Sigma：http://www.sigma.com

氨基硅烷（硅烷化）

氨基硅烷是替代聚 L-赖氨酸常用的包被。制备了各种氨基修饰的硅烷与玻璃的硅烷醇表面反应。在中性 pH 下，硅烷上带正电的氨基与带负电的 DNA 磷酸二酯骨架之间的静电互相作用将 DNA 结合到基底（Halliwell and Cass 2001）。UV 照射通过产生自由基介导的胸腺嘧啶苷残基与烷基胺上的碳原子之间的耦联，加强和稳定了这种相互作用。氨基硅烷包被几乎对高温不敏感，并且不会出现与聚 L-赖氨酸同样的脱落问题。如果出现不均一的包被，或者未使用的包被表面没有完全封闭，玻片部分位置将发出红色的荧光背景信号。廉价的、用于组织学应用的氨基硅烷包被的玻片，可能会有不均匀的包被，导致高的背景信号及假象。为微阵列设计的高质量的氨基硅烷包被可以购买

到，并且通常产生较好的结果。应用该类型包被制备的 DNA 微阵列的样点大小在不同的供应商之间变化很大，这表明在不同的商业配方之间，疏水性的程度不同。该包被不存在保存期的问题。

供应商：

Corning Microarray Technology（CMT）（http://www.corning.com）：γ-氨丙基硅烷（CMT-GAPSⅡ）

Telechem/ArrayIt. com（http://www.arrayit.com）：SuperAmine

Clontech：http://www.clontech.com/ TypeⅠ，TypeⅡ. TypeⅠ氨基硅烷化玻片推荐用于大于 200 bp 的 DNA，TypeⅡ玻片推荐用于较短的 DNA 探针（＞50 bp）。

Eppendorf（www.brinkman.com）：Creative Chip Oligo 和 Creative Chip PCR 载玻片。

醛（甲硅烷基化）

DNA 分子结合到中性疏水包被中的醛基上，通过与化学修饰的 DNA 上 5′氨基手臂分子共价结合，或者是共价连接到 DNA 链中 G、C 和 A 残基的芳香胺。随后包被中未反应的醛基与硼氢化钠反应还原为伯醇。制备具有 5′端氨基用于点样的 DNA 可以使用带有氨基的 PCR 引物（见第 1 章，表 1-1），或者在长链寡核苷酸的合成中使用修饰的磷酸酰胺（见第 1 章，表 1-6）。与直接和 DNA 分子中的氨基反应相比较，通过 5′端氨基连接到基底的反应效率更高。与有更多机会和 DNA 链中碱基反应的 cDNA 片段相比，该考虑对于寡核苷酸特别重要。醛基修饰的基底具有高的结合容量，提供了牢固的接合，显示较低的荧光背景，并且在与几种常用基底面对面的直接比较中，至少有一次更显示出优势（Zammatteo et al. 2000）。然而，醛基玻片在空气中储存时间过长会导致一部分醛基由于被氧化形成羧酸而丢失。

供应商：

Telechem/ArrayIt. com（http://www.arrayit.com）：SuperAldehyde

Genpak（http://www.genpakdna.com/slid2.shtml）：Silylated

Quantifoil（http://www.quantifoil.com）QMT aldehyde

活化的聚合物包被

这类包被由长链的亲水聚合物形成，含有氨基反应基团，可以形成链间交联和与玻片表面结合。聚合物包被可以与带有末端氨基的 PCR 产物和寡核苷酸共价结合，固定 DNA 并使其离开玻片表面，从而使它更容易杂交，同时避免了在寡核苷酸上接上 poly（dT）或 PEG 手臂分子的需要。聚合物的亲水特性减少了玻片表面的非特异性结合，给出了低的背景。带有末端氨基的 DNA 样品重悬在 150 mmol/L 磷酸钠溶液中用于点样，点样后的玻片在潮湿的盒中反应过夜，然后按照方案 2 所述内容进行清洗和封闭。活化的玻片保存期可以超过一年。

供应商：Motorola Life Science（http://www.motorola.com/lifescience）

其他方法

将 5′二硫化物修饰的寡核苷酸结合到巯基硅烷包被的表面（Rogers et al. 1999）代表了一种在概念上与使用氨基修饰过的 DNA 和醛基表面类似的方法。

参考文献

Call D.R., Chandler D.P., and Brockman F. 2001. Fabrication of DNA microarrays using unmodified oligonucleotide probes. *BioTechniques* **30:** 368–376.

Diehl F., Grahlmann S., Beier M., and Hoheisel J.D. 2001. Manufacturing DNA microarrays of high spot homogeneity and reduced background signal. *Nucleic Acids Res.* **29:** E38.

Halliwell C.M. and Cass A.E. 2001. A factorial analysis of silanization conditions for the immobilization of oligonucleotides on glass surfaces. *Anal. Chem.* **73:** 2476–2483.

Hegde P., Qi R., Abernathy K., Gay C., Dharap S., Gaspard R., Hughes J.E., Snesrud E., Lee N., and Quackenbush J. 2000. A concise guide to cDNA microarray analysis. *BioTechniques* **29:** 548–556.

Rogers Y.H., Jiang-Baucom P., Huang Z.J., Bogdanov V., Anderson S., and Boyce-Jacino M.T. 1999. Immobilization of oligonucleotides onto a glass support via disulfide bonds: A method for preparation of DNA microarrays. *Anal. Biochem.* **266:** 23–30.

Schena M., Shalon D., Davis R.W., and Brown P.O. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270:** 467–470.

Schena M., Shalon D., Heller R., Chai A., Brown P.O., and Davis R.W. 1996. Parallel human genome analysis: Microarray-based expression monitoring of 1000 genes. *Proc. Natl. Acad. Sci.* **93:** 10614–10619.

Yue H., Eastman P.S., Wang B.B., Minor J., Doctolero M.H., Nuttall R.L., Stack R., Becker J.W., Montgomery J.R., Vainer M., and Johnston R. 2001. An evaluation of the performance of cDNA microarrays for detecting changes in global mRNA expression. *Nucleic Acids Res.* **29:** E41.

Zammatteo N., Jeanmart L., Hamels S., Courtois S., Louette P., Hevesi L., and Remacle J. 2000. Comparison between different strategies of covalent attachment of DNA to glass surfaces to build DNA microarrays. *Anal. Biochem.* **280:** 143–150.

因特网资源

http://abrf.org/ABRF/Research Grills G. et al. 2000/2001. ABRF Microarray Research Group study: A current profile of microarray laboratories.

http://affymetrix.com Affymetrix homepage.

http://arrayit.com TeleChem International, Inc. homepage.

http://cmgm.stanford.edu/pbrown/mguide/index.html The Brown lab's complete guide to microarraying for the molecular biologist.

http://majerprecision.com Majer Precision Engineering, Inc. homepage.

http://www.Biorobotics.com BioRobotics Ltd. homepage.

http://www.cartesian.com Cartesian homepage.

http://www.clontech.com Clontech homepage.

http://www.corning.com/CMT Corning homepage.

http://www.genetix.com Genetix Ltd. homepage.

http://www.genpakdna.com/slid2.shtml GenPak's aldehyde coated (silyated) slides.

http://www.menzel.de/menzel/index.html Menzel-Gläser homepage.

http://www.motorola.com/lifesciences Motorola Life Sciences homepage.

http://www.mscdirect.com MSC Industrial Supply Company homepage.

http://www.packardbiochip.com Packard BioScience microarray and biochip technologies.

http://www.probes.com Molecular Probes homepage.

http://www.quantifoil.com Quantifoil zur deutschen homepage.

http://www.sigma.com Sigma-Aldrich homepage.

http://www.virtekbiotech.com Virtek Biotechnology homepage.

101

第3章 RNA的表达分析

导言

许多微阵列项目的目的，是要确定不同类型的细胞或不同生长条件下同种细胞的复杂 RNA 样品中，基因表达丰度的不同。

过去几年中，发展了一些在制备的总 RNA 或 Poly（A）$^+$ RNA 中测量目标信使 RNA（mRNA）丰度的方法。这些方法中最常用的——Northern 印迹法和核糖核酸酶保护法直到今天仍然通用。这两种方法具有同样的局限性：它们使用一种标记的、特异的核酸探针去检测一种（或者最多几种）制备物中特异性的 mRNA；通过间接比较从目标 mRNA 得到的信号和从一个标准的管家基因或是已知浓度的人工目标 RNA（artificial target RNA）获得的信号来实现 mRNA 的定量（关于进一步的讨论，请见第 7 章 Northern 杂交中的导言；Sambrook and Russell 2001）。类似的不足也适用于一些其他技术，如定量反转录 PCR（RT-PCR）。尽管它们有这些局限性，然而分析单独的或者多种特定 mRNA 丰度的方法可以敏锐洞察许多“单基因”疾病的生物学和病理学。但是，理解、分类和处理更复杂的疾病，如癌症，则需要全局性地分析基因的表达。

103

现代微阵列的许多原理建立在 20 世纪 80 年代末和 90 年代初，当时点样在膜上的

克隆 cDNA 用来和复杂的靶序列杂交，以在较大的动态范围内对 mRNA 的表达差异进行定量（Freimert et al. 1989；Lennon and Lehrach 1991；Gress et al. 1992；Takahashi and Ko 1994；Maier et al. 1994；见第 4 章）。在 20 世纪 90 年代中期有了一个重大的突破，当时 Pat Brown，Ron Davis 和他们的同事报道了一种双色标记的、内在的比较杂交技术（internally comparative hybridization technique）来检测以高密度方式自动点样在固相基底上的 cDNA（Schena et al. 1995，1996；DeRisi et al. 1996）。这些研究使得可以同时筛选数千种 mRNA 的相对丰度的 DNA 微阵列诞生。图 3-1 中描述了各种 cDNA 微阵列的检测方案和应用。

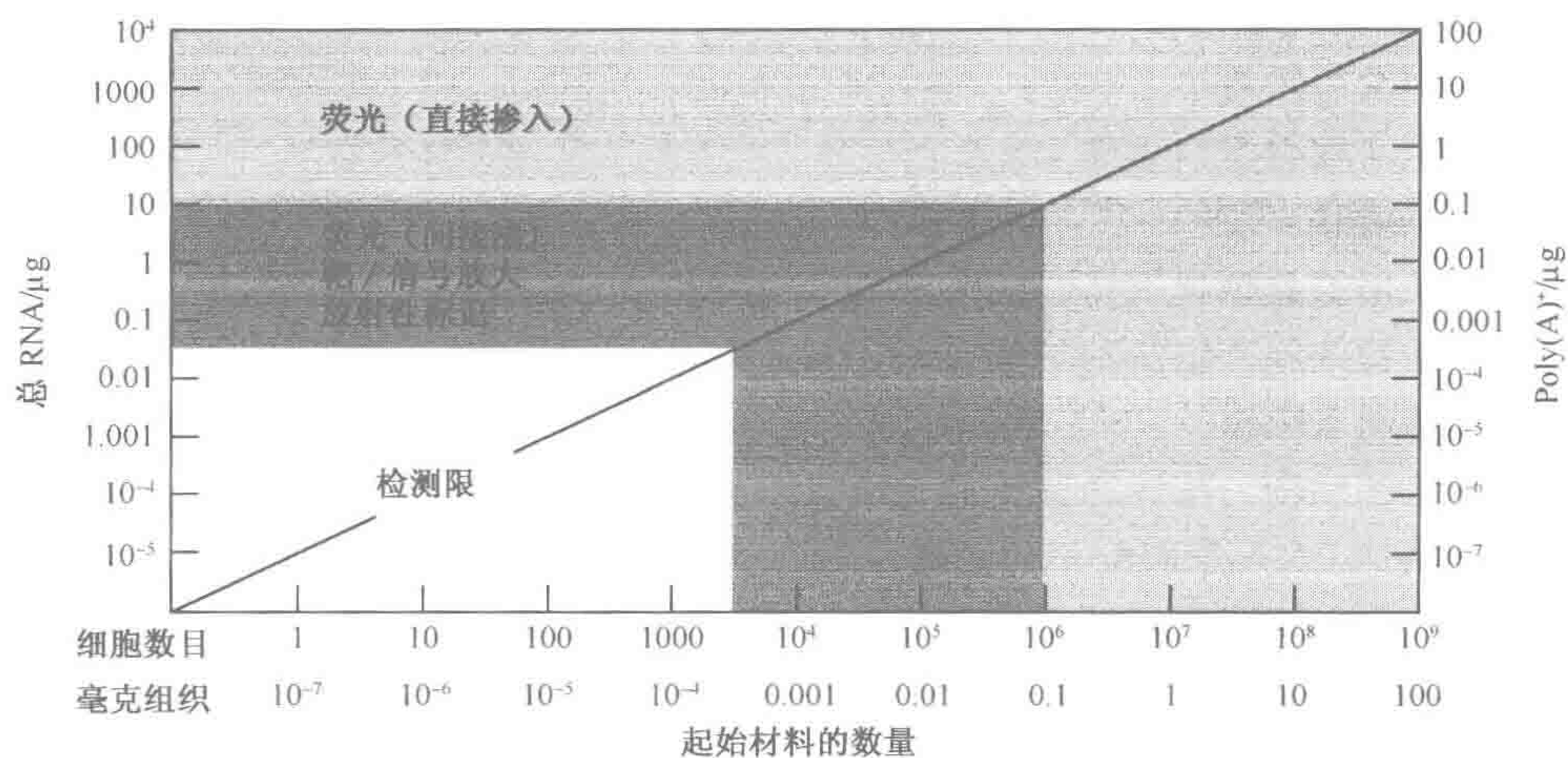


图 3-1 cDNA 微阵列的检测方案和应用

基因表达量的变化可以使用多种方案进行检测，但检测限不同。在 cDNA 靶序列中直接掺入带荧光的核苷能够用来从 10 μg 或更多的总 RNA 中检测表达谱。另一方面，间接荧光、靶序列与信号的放大以及放射性标记，能够检测少至 50 ng 的总 RNA 的表达谱。这个检测限可以研究来自大量生物来源的，包括细胞培养物、临床活检样品（包含尸检材料）和组织学标本的表达谱。（经许可重新绘制，来自 Duggan et al. 1999）

这一部分讲述了如何用 DNA 微阵列来获得细胞和组织中 mRNA 的全局性表达谱特征。PART Ⅰ（方案 1～方案 12）描述了从多种来源制备适合于表达分析的 RNA 的方法。方案 11 概述了如何在体外扩增少量的 RNA 来提供足够的模板进行直接或间接标记（方案 13 和方案 14）。PART Ⅱ（方案 13～方案 20）继续描述了将标记的 RNA 与 DNA 微阵列杂交的技术（方案 20），并提供了广泛的疑难问题解答指南。

104

PART Ⅰ RNA 分离

导言

表达分析的成功与 RNA 的质量紧密相连（Schena and Davis 2000）。被盐、多糖、DNA、蛋白质或脂类污染的 RNA 标记效率低并且杂交时会产生高的背景。如果 RNA 部分降解，标记将偏向 3′端的序列，或者是抗 RNase 降解的序列。这可能会破坏通过与 DNA 微阵列杂交来检测的各种 RNA 的相对比例。因此，精确和灵敏地估计 mRNA 的相对或绝对含量，取决于研究人员提取纯化未损伤的 RNA 制备物的能力。

一旦组织从动物体切除，mRNA 合成的整个速率就会降低，而降解的速率增加。成功分离用于表达分析的 RNA 的关键是细心地制备和收集细胞和组织并快速将它们裂解。使用可以抑制 RNase 的方式来制备和裂解组织和细胞，将会显著提高成功的可能性。表 3-1 概述了在抽提和纯化总 RNA 时常用的收集和裂解组织、细胞的方法。这些方法在方案 1～方案 10 中进行了详细的描述，介绍了从动物组织和细胞、植物及酵母中 RNA 的分离。更进一步的信息可以从**基于胍盐的 RNA 抽提**信息栏和表 3-2 中得到。方案 11 描述了从少量材料或单个细胞中分离 RNA 的方法；使用激光切割技术从复杂组织的少量细胞中回收 RNA 的其他方法可以在手册的第 5 章中找到。最后，在设计包含多种不同的 RNA 制备物的研究中，使用一个通用的 RNA 参考集，在减少任务的复杂性（和艰巨性）方面提供了独特的有利条件。方案 12 中描述了 RNA 通用参考集的制备和使用。

表 3-1　制备用于 RNA 分离的生物样品

制　备	保存条件[a]	裂解方法
人和动物组织样品（方案 1，2，5） 将人的活检或者手术标本迅速在冰上冷藏，尽快送给病理学家。征得患者同意后，病理学家将提供不需要用于诊断的组织。大多数病理学家宁愿切割新鲜未处理的组织，而有些同意在切割和固定用于组织学检测之前先把组织保存在 RNA*later*[b]（Ambion 7020）中。RNA*later* 为进行微阵列分析时能保持 RNA 的完整性，且不会危害用于病理诊断的组织学和免疫组织化学技术（Florell et al. 2001）。在任何情况下，组织取样都不能影响到诊断；所有决定，包括组织处理和用于研究的组织数量必须由病理学家作出	采集后尽快处理用于分离 RNA 的样品。尽快将从动物切下或者病理学家提供的组织切成在任何方向上厚度不大于 0.5 cm 的碎片；将组织碎片放入 5 倍体积的 RNA*later*。将样品在室温下保存 5 min，然后转入 4℃ 下保存最多一个月或在 −20℃ 保存更长的时间 如果长期保存，按上述方法用 RNA*later* 处理样品，然后移至冻存管中。在液氮中将组织速冻。冷冻管置于 −80℃ 下长期保存	加入裂解试剂之前，用液氮冷冻的研钵和研杵将坚韧的（皮肤，骨）或纤维状的（肌肉）组织研磨成粉末。将粉末加入到含有单相裂解试剂[c]的聚丙烯试管中 通常，软组织如脾和肝采用单相裂解试剂在转子—定子匀浆器（如 Polytron）中裂解（见方案 1）。更多详细内容，请见**动物细胞和组织的匀浆**信息栏
培养细胞样品（方案 1，2，5） 尽可能地使细胞生长标准化。为了方便不同实验之间结果的比较，尽可能地精确的重复生长条件。使用密度相同，生长于同样的培养瓶和同一批培养基中，处在同样的温度/湿度条件下的培养物。尽量减少孵育和冷冻/裂解细胞之间的时间。确保培养物中没有支原体	刮取单层培养细胞，通过离心收集或者在原位裂解细胞。离心收集悬浮培养的细胞 如果进行 RNA 抽提前要保存细胞，则要将团块在少量（约 200 μl）PBS 中重悬，并冷冻在试管底部以便融化时迅速暴露到裂解缓冲液中（将细胞在 PBS 中重悬成匀浆状，向侧面倾斜试管并旋转浸入液氮）。作为替代，加入 5～10 倍体积的 RNA*later*，可以将细胞在 4℃ 下保存最多一个月，或者在 −20℃ 长期保存	将洗涤过的细胞团块重悬于单相裂解液中 冷冻单层生长的细胞，然后在培养板上直接裂解。将细胞裂解物转移到聚丙烯试管中，剧烈涡旋振荡或者反复上下吸打，使细胞彻底瓦解

制　备	保存条件[a]	裂解方法
植物样品（方案3，4，7） 要小心生长室条件，光照时间、温度和 CO_2 分压是样品间改变显而易见的可能来源。植物生长和处理中常常忽略的一些方面（受精条件、植物年龄、生长培养基、缓冲液、激素和其他添加剂）也可能是改变的来源。为避免出现问题，所有样品的光强和光谱、一天中组织处理的时间和采集都应该相同。尽可能减少植物多糖和复杂有机化合物积累。这些物质是植物核酸制备中常见而顽固的污染物，会抑制多种酶促反应，包括热稳定性聚合酶和反转录酶	在 RNA*later* 中洗涤新鲜的植物组织（可选的），然后立即用 TRIzol 裂解。作为替代，将洗涤过的组织在－80℃无限期地冷冻。液氮下用研钵和研杵研磨冷冻的组织，并以粉末保存	使用转子—定子匀浆器，搅拌研磨器（如 MM300 型，QIAGEN）或者 Polytron，在裂解缓冲液中匀浆裂解来自植物，如 *Arabidopsis* 和 *Nicotiana* 的新鲜软组织。其他来自坚硬的木本植物的组织可能需要冷冻并在液氮中研磨。将裂解物通过一根粉碎离心柱（如 QIAshredder，QIAGEN），可以降低用单相裂解试剂处理的植物裂解物的黏度
酵母样品（方案9） 自养生物，如酵母，表达分析的目的是对特异性刺激或遗传改变导致的基因转录物丰度的变化进行分类（Schena et al. 1995；DeRisi et al. 1997；Gasch et al. 2000）。因此，用于 DNA 微阵列分析的酵母培养物的生长条件主要取决于环境刺激的性质和酵母菌株的遗传组成。然而，大多数实验包括在标准条件下，如添加 0%～25%（*m*/*V*）葡萄糖的 YPD 培养基（Sherman 1991）对照培养物的生长。培养物通常以不超过培养瓶体积的 25% 为宜。在 30℃下以 250～300 r/min 振荡培养	通常不需要使用 RNA*later*。生长期为早至中对数期（OD_{600} = 0.2～0.6）的细胞可以离心收集并重悬于 50 ml 适当的裂解缓冲液中，在液氮中快速冷冻，保存于－80℃直到 RNA 抽提	加入酸性酚裂解酵母细胞（见方案 9）。不能通过冷冻或其他机械方法破坏的细胞壁结构用酶消化。然后在 65℃用水饱和的酚抽提 RNA，再用异丙醇或乙醇沉淀。作为替代，使用试剂盒如 QIAGEN RNeasy（见方案 1）

对于一些研究人员，一种经济的替代方法是从合适的供应商处购买非试剂盒装的单相裂解液和树脂：TRIzol 试剂（Invitrogen/GIBCO），TRI 试剂（Molecular Research Center），RNA STAT60（Tel-Test，Friendswood，Texas），TOTALLY RNA（Ambion），Isogen（Nippon Gene Toyama，Japan）。TRIzol 是目前纯化 RNA 用于 DNA 微阵列表达分析中使用最广泛的。TRIzol 是苯酚、胍、硫氰酸铵、醋酸钠、酚红和甘油组成的单相溶液（配方见 http://afgc.stanford.edu/afgc_html/AFGCProtocols-fev2001.pdf）。硫氰酸铵的离液作用溶解细胞组分并使蛋白质变性，同时保持核酸的完整。向裂解物加入氯仿使溶液分成两相。下层红色有机相含苯酚和氯仿。在高盐浓度下用异丙醇沉淀可以回收水相中的 RNA。然后进行层析，层析可以去除相分离残留的 DNA 和盐，从而纯化 RNA。关于详细内容，请见**基于胍盐的 RNA 抽提**信息栏。

a 维护包含样品所有相关信息，包括采样和保存之间的时间长短和移入 RNA*later* 或－80℃之前组织在每个温度下的停留时间的数据库。以防乙醇的记号笔或 Tough Tags、Tough Spots（Diversified Biotech，Boston Massachusetts）标记保存管。

b RNA*later* 是一种液态无毒的组织和细胞保存试剂。根据供应商（Ambion）的资料，RNA*later* 在数分钟内渗透组织（任何方向上厚度小于 0.5 cm）并失活 RNase。组织样品能在 5 倍体积的 RNA*later* 中 4℃下保存最多一个月，或在－20℃长期保存。分离 RNA 时，从 RNA*later* 中离心回收细胞/组织，并用单相裂解试剂裂解。

c 根据最新统计，有 55 种商品化试剂盒用于从细胞中分离总 DNA（http://www.biocompare.com）。它们中的大多数含有各种裂解试剂的组合——不一定是单相的——同时还有用于裂解后纯化 RNA 的树脂和溶液［关于裂解试剂组成的综述，见 Defrancesco（1998）］。虽然这些试剂盒价格昂贵，但仍然有充分的理由使用它们制备微阵列实验需要的 RNA。这些试剂盒通常使用简单、有效，且最重要的是可重复。比较一系列样品的表达谱时，一个重要目标是应该尽量减少组织和试剂操作及处理中的变化。

表 3-2 基于胍盐的 RNA 抽提中疑难问题解答指南

问 题	可能的原因	建议的解决方法
RNA 降解	细胞或组织采集与裂解间隔时间过长	使采集和裂解之间的时间间隔最小。采集后将细胞和组织立即冷冻。用 RNA*later* 洗涤和保存细胞及组织。如果细胞是单层生长，直接向培养板中加入单相裂解试剂（如 TRIzol）
	用于失活/纯化 RNA 的试剂量过多（常见问题）	为组织/细胞使用推荐比例的试剂。不能少加！通常每 0.05 g 组织或每个 100 mm 直径培养皿使用 1 ml TRIzol（表 3-3）
	来自界面的物质污染水相	小心移液。不要试图获取最后一滴水相
	移液时，高浓度的高分子质量 DNA 引起的黏滞破坏了界面	使用适量的 Trizol、匀浆或将溶液通过细针剪碎 DNA
	水溶液中 RNA 不能冷冻且不受保护	将温度大于 0℃ 的时间减至最少。RNA 将保存在乙醇中（见 **RNA 的保存和回收**信息栏）
	缓冲液或设备被 RNase 污染	按照 **RNA 操作的特别注意事项**信息栏中给出的建议
	初始组织富含核酸酶（脾、胰、胸腺）	减少取样和处理之间的时间。将组织冷冻在液氮中并在冷冻状态下研磨成粉末。研磨结束后立即加入 TRIzol
RNA 产量低	组织和细胞在 TRIzol 试剂中未能完全匀浆	有效地匀浆总是很重要，并且在处理纤维性的组织如心脏和骨骼肌时是关键所在。将要匀浆的组织在液氮中冷冻，并在冷冻状态下研磨成粉末。确保研磨结束后没有可见的颗粒物质存在。加入 TRIzol 匀浆后在室温下温育制备物 5 min。如果还有纤维状物质，将裂解物用 Polytron 匀浆器匀浆 30 s。如果有必要，继续处理前通过离心去除所有残留的不溶物质
	用 TRIzol 孵育时，细胞不能全部从培养瓶/皿中移走	向培养瓶/皿中加入 TRIzol，前后倾斜溶液。竖直放置培养瓶/皿 3～5 min，用橡胶刮勺将细胞刮到培养瓶/皿的一侧。继续竖直放置培养瓶/皿 3～5 min，然后用吸管吸取裂解剧烈地清洗培养瓶/皿的表面。将裂解物移入试管，在室温下孵育 5 min
	RNA 不能完全溶解	用乙醇或异丙醇沉淀后，不要让 RNA 团块干燥。不要真空干燥团块。完全脱水的 RNA 十分难溶，尝试在 55℃ 无 RNase 的水中溶解
	细胞/组织的转录活性比预期要低	原代细胞培养物、成片的单层不分裂细胞和脂肪含量高的组织中 RNA 的量都低
低 $A_{260/280}$ 比率（<1.70）	制备物中残留有机溶剂（苯酚、氯仿）	确保相分离完全。如有疑问，进行离心。小心操作，不要把有机相带入 RNA 样品 用 2.5 倍体积乙醇重悬 RNA
	样品中加入的 TRIzol 量不足	一般情况下，每 0.05 g 组织或每个 100 mm 直径培养皿中加入 1 ml TRIzol（表 3-3）
	溶液 pH 为酸性	用 TE 代替水来溶解 RNA
	制备物中残留有异硫氰酸胍	加入 0.1 体积的 2 mol/L 醋酸钠溶液（pH 4.0）。加入 1 体积的异丙醇。将溶液在 −20℃ 保存 30 min，离心回收 RNA。用 70% 乙醇仔细洗涤团块，并让其在空气中干燥。团块未干透时，用 DEPC 处理过的 H_2O 溶解，并用乙醇沉淀 RNA（3 倍于 RNA 溶液体积的乙醇）。离心回收 RNA
	A_{260} 或 A_{280} 在比色计的线性范围之外	测量 RNA 制备物稀释样品的 OD

续表

问　题	可能的原因	建议的解决方法
RNA 制备物中含有一些 DNA	部分界面物质随水相转移	回收水相时确保不带入界面上的物质
	样品中加入的 TRIzol 量不足	一般情况下，每 0.05 g 组织或每个 100 mm 直径培养皿中加入 1 ml TRIzol（表 3-3）
	组织的溶解不完全	匀浆结束后确保没有可见的粒状物质残留。匀浆后在室温下将制备物温育 5 min。如有必要，继续处理前进行离心以去除不溶的物质
	裂解试剂：组织的比率过低。脾和胸腺类组织中含有大量的 DNA，必须增加 TRIzol 试剂用量以使裂解液的黏度减少到容易处理的水平。如果问题依然存在，进行匀浆或将裂解物流过细针来剪切 DNA	增加 TRIzol：组织的比率；延长高速离心的时间以分离水相和有机相；重复抽提以确保 DNA 进入酸性苯酚相；回收水相时避免带入界面；在往酸性苯酚相中加入异丙醇后若产生白色沉淀，则表明 DNA 污染仍然存在，这种情况下，回收沉淀，用不含 RNase 的水重悬，然后用氯仿重新抽提数次
含 RNA 的水相带有颜色（微黄褐色或粉红色）	在离心过程中脂肪滴未浮到水相上面。试剂中的酚红集中到脂肪滴中	加入氯仿之前离心制备物，并用连接到真空管的 Pasteur 吸管除去表层微黄色的脂肪
	如果样品中血液含量高，RNA 制备物中可能残留有血红蛋白	裂解时增加裂解试剂对样品的比率
	加入异丙醇后水相变为黄色	使用一瓶新的异丙醇
RNA 虽然完整，但是反转录酶的低效模板	RNA 制备物被多糖和（或）多酚污染。这个问题常见于植物的 RNA 制备物	分离 RNA 前用聚乙烯吡咯烷酮（PVP）处理植物组织裂解物（见方案 4 中分离 RNA 的松树方法）

本表中的一些材料从 Invitrogen、Promega Ambion 网站复制。

预防 RNase 的降解来自动物细胞

核糖残基 2′和 3′位置上的反应性羟基使 RNA 易受到 RNase 的攻击。这种普遍存在并且稳定的酶在细胞裂解时释放，并且存在于皮肤上。尽管许多实验因为 RNase 的污染而遭到破坏，但通过警惕和常识完全可以避免这个问题。常识性的预防措施列举在 **RNA 操作的特别注意事项**和**抑制 RNase** 两栏信息栏中。

检查来自动物细胞的 RNA 制备物的完整性

在溴乙啶存在条件下对 RNA 电泳后，28S 和 18S 的核糖体 RNA（rRNA）在紫外线照射下应该清晰可见，同时还有一些由转移 RNA（tRNA）和 5S 核糖体 RNA 组成的更加扩散且快速迁移的条带。如果 RNA 制备物未被降解，28S 核糖体 RNA 条带的染色强度应该是 18S 条带的大约两倍，并且这两个条带应观察不到拖尾现象。除非凝胶超载，否则不能观察到 mRNA（图 3-27）。下面是三种能够检测 RNA 完整性的其他方法。

- **用 oligo (dT) 作为引物分析合成的 cDNA 的长度** 在试验性的第一链 cDNA 的反应中合成具有放射性的互补 cDNA。从哺乳动物 mRNA 合成的 cDNA 在电泳后应该是从约 600 bp 到大于 5 kb 连续扩散分布的。主要的放射活性位于 1.5～2 kb，并且没有 cDNA 条带出现，除非 mRNA 从高度分化的、大量表达一套特定蛋白质的细胞（如网状细胞和 B 淋巴细胞）制备。
- **在试验性 Northern 杂交中用放射性标记的 oligo (dT) 作为探针** (Fornace and Mitchell 1986；Hollander and Rornace 1990)。Poly (A)$^+$RNA 的放射自显影图谱从约 600 bp 到大于 5 kb 连续扩散分布。主要的放射活性位于 1.5～2 kb，并且没有 cDNA 条带出现，除非 mRNA 从表达大量特定 mRNA 的细胞制备。
- **用 Northern 杂交检测已知长度的目标 mRNA** 例如，一个来自于管家基因表达的 mRNA 如葡萄糖-6-磷酸脱氢酶 (GAPDH) mRNA，在大多数类型哺乳动物细胞中的长度约为 1.3 kb。条带在低分子质量区域的显著扩散通常表明发生了总 mRNA 的显著降解。另一方面，条带模糊不清是琼脂糖凝胶系统出现问题的征兆（如在乙二醛中存在离子）。关于进一步的详细信息请参见分子克隆：实验手册的第 7 章 (Sambrook and Russel 2001)。

DNA 的污染

制备的总 RNA 在凝胶电泳后，加样孔附近存在被溴化乙啶染色的物质是基因组 DNA 仍然存在的确定迹象。一个更严格的测试是用一小部分制备的 RNA 作为模板，用对应于一个内含子中已知距离的序列的寡核苷酸作为引物进行 PCR。如果反应产生的产物的长度与预测长度相等，那么制备物受到 DNA 的污染。在这种情况下，总 RNA 制备物应该用不含 RNase 的 DNase 制备物处理（如 QIANGEN 不含 RNase 的 DNase 和 RNeasy/QIAamp 柱一道使用)；作为替代，可以从制备的总 RNA 中分离 poly (A)$^+$RNA。用氯仿抽提细胞裂解物的过程中，当有机相和水相界面的物质污染水相时，会发生 DNA 的污染。

方案 1

110

来自哺乳动物组织和细胞的总 RNA 的纯化

Jun S. Wei and Javed Khan (先进技术中心，儿科肿瘤学部门，国家癌症研究所，NIH，盖瑟斯堡，马里兰，20877)

在有机溶剂存在情况下进行离液裂解是从绝大多数类型哺乳动物的组织和培养细胞中提取 RNA 的首选方法（参看**基于胍盐的 RNA 抽提**信息栏)。本方案中，在 TRIzol 裂解后进行相分离除去 DNA 和蛋白质。随后通过与层析介质（这里是 RNeasy 柱）结合、清洗和洗脱，从离液溶液中纯化和回收 RNA (图 3-2)。这个方法有许多优点。由于省略了早期的浓缩和沉淀步骤，避免了 RNA 和多糖、蛋白聚糖之间形成复合物。这些复合物阻止了沉淀后的 RNA 溶解，并抑制反转录反应。由于同时使用了相分离和层

析，它们中任何一种方法都建立在不同的物理基础上来分离分子，这要比单独使用任何一种方法具有更好的纯化结果。信号强度的直接对比表明，从总 RNA 制备的 cDNA 的信号强度要比从 poly (A)$^+$ RNA 获得的信号强度高大约 3 倍 (Wildsmith et al. 2001)。这可能是由于 poly (A)$^+$ RNA 分离时 mRNA 的潜在损失，因此与使用总 RNA 相比没有优点。

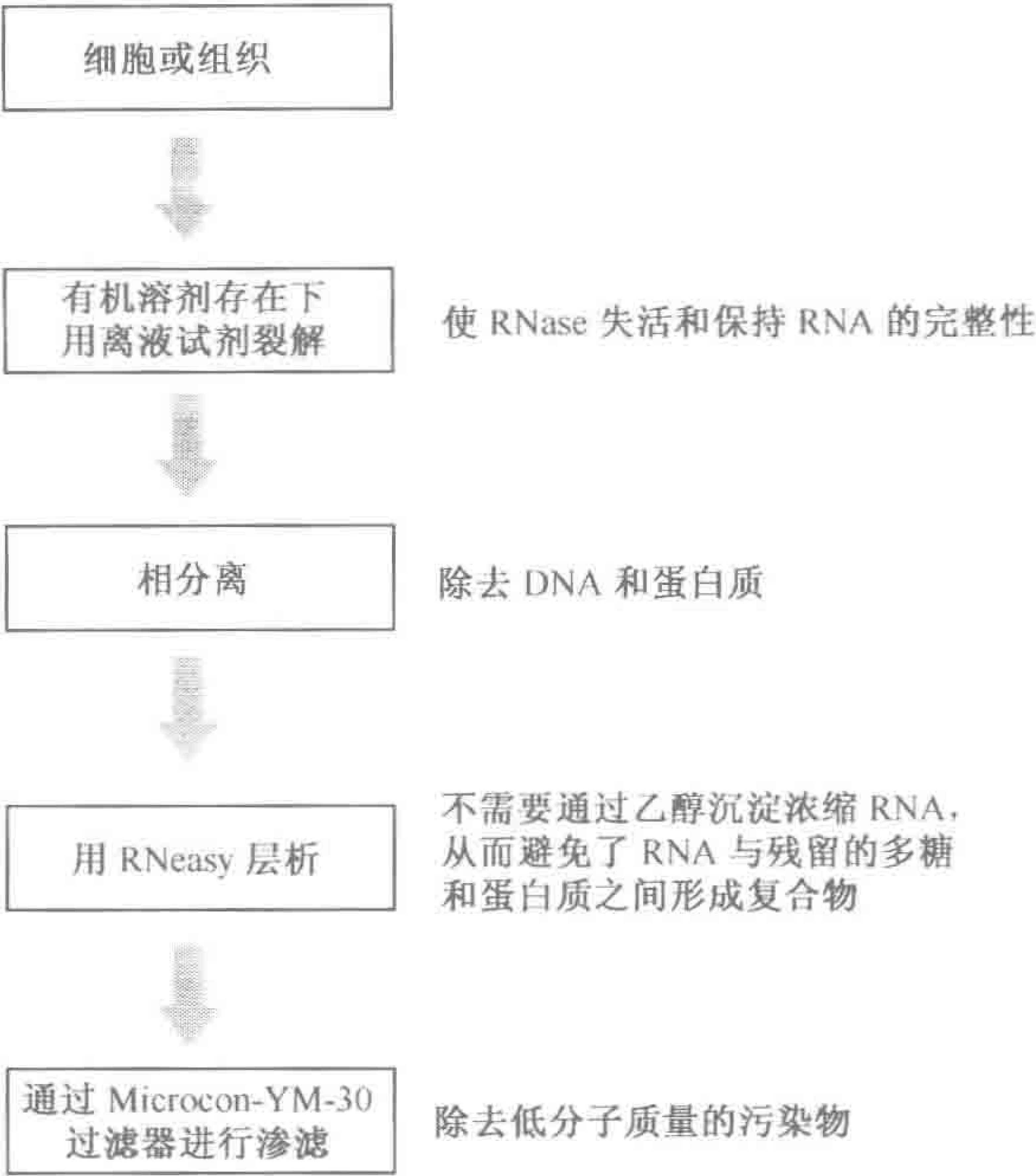


图 3-2　分离总 RNA 的程序

流程图中显示了从细胞和组织中分离总 RNA 的优化程序，制备物用于荧光标记

(关于实验细节，请见下文)

尽管这里描述的技术已经被成功应用于许多组织和细胞，但是在从富含甘油三酯的脂肪组织中抽提 RNA 时，不推荐应用这种方法。从这些脂肪丰富的原材料中制备 RNA 的最佳方法是 Tavanger 等 (1990) 描述的经过修改的 Chirgwin (1979) 方法。也见本部分的方案 8。

在本方案最后的信息栏中推荐了从少量的组织中成功抽提 RNA 的方法。同样在方案最后有一个替代方案**从 OCT 包埋的病理样品中提取 RNA**。 111

112 材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

重要事项：本方案中所有试剂使用经 DEPC 处理的 H_2O 配制 (见 **RNA 操作的特别注意事项**信息栏)。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

氯仿〈!〉

DEPC 处理的 H_2O〈!〉（见**抑制 RNase** 的信息栏）

乙醇（无水）

据报道，有些批次的乙醇（100%或 200 标准强度）中含有会使微阵列试验中产生很深荧光背景的污染物质，95%的乙醇不存在这个问题。为避免这个问题，可以用 95%的乙醇或者已证明不会产生高荧光背景的无水乙醇。Warner-Graham 公司（Cockeysville，Maryland）出售的无水乙醇（USP 级乙醇，64-17-5）效果很好。

单相裂解溶液（15596，TRIzol，GIBCO）（表 3-3 和表 3-4）

磷酸盐缓冲液（PBS），不含 Ca^{2+} 和 Mg^{2+}（14190-136，GIBCO）

RNA*later* 稳定剂（221462，QIAGEN）

RNase*Zap*（9780，Ambion）

醋酸钠（3 mol/L，pH 5.2）

表 3-3 从单层细胞中分离 RNA 需要的单相裂解试剂的量

培养皿	表面积/cm^2	标准单相裂解试剂的最小体积（如 TRI 试剂）/ml	浓缩的单相裂解试剂的最小体积（如 TRI 试剂 LS）/ml
6 孔培养皿每个孔	10.00	0.80	0.60
10 cm 培养皿	78.50	7.80	2.75
T75 长颈瓶	75.00	5.25	2.65
T750 长颈瓶	150.00	10.50	5.30

单相裂解试剂的量是根据培养皿的面积而不是细胞的密度确定的。表中列出的试剂体积足够完全覆盖培养板或长颈瓶，并可以克服残余的培养基或磷酸盐缓冲液的缓冲效应。

一些制造商销售的裂解试剂浓度不同，如标准的 TRIzol 和更为浓缩的 TRIzol LS。表中 TRI 试剂的数据网站，Molecular Research Center Inc. 的单相裂解试剂的形式是 TRI 试剂（等同于 TRIzol）和 TRI 试剂 LS（等同于 TRIzol LS）（http://www.mrcgene.com/tbl-rna.htm）。

表 3-4 单相裂解试剂

试　剂	供应商
TRIzol	Invitrogen（Carlsbad，California）
TRI Reagent	Molecular Research Center（Cincinnati，Ohio）
Trisolv	Biotecx Laboratories（Houston，Texas）
RNA STAT60	Tel-Test（Friendswood，Texas）
Isogen	Nippon Gene（Toyama，Japan）

使用商品化试剂时应查看制造商的说明书。

细胞和组织

哺乳动物细胞的单层或悬浮培养物或新鲜组织

专用设备

各种大小带滤芯的微量吸头

离心管（15 ml 带旋盖；60-540，Sarstedt，Newton，North Carolina）

医用台式离心机

匀浆器

配备 Probe 9100071（Brinkmann 2713 700-8）的 Polytron PT1200 匀浆器（Kinematica AG，Switzerland）或相当的仪器被用来捣碎细胞和组织。使用前，清洁 Polyton 使其没有 RNase。拆开探针，先用水洗涤部件，然后用无水乙醇清洗，最后用可降解 RNA 的 RNaseZap 处理。组装探针并在水中匀浆。关于组织匀浆使用的各种匀浆器的讨论，见**动物细胞和组织的匀浆**信息栏。

Microcon YM-30（42410，Millipore）

研钵和研杵

捣碎大量的（>500 mg）的软组织和任何数量坚硬的纤维性组织时所必需的（见步骤 1）（12-061-5B，Fisher Scientific，Pittsburgh Pennsylvania）.

冷冻离心机（Sorvall RT 6000/H1000 转子或相当的）

RNeasy RNA 分离试剂盒（Midi kit 75144；Maxi kit 75162，QIAGEN）

SpeedVac

方法

1. 准备分离 RNA 的组织或细胞。

对于组织：

只要可能，组织匀浆前将组织保存在－80℃冷冻，或在液氮或 RNA*later* 中直到匀浆。

a. 组织称重。

b. 将大约 100 mg 的冰冻组织放入一个含有 4 ml TRIzol 的 50 ml 带旋盖的聚丙烯离心管中。

c. 用 Polytron 匀浆器匀浆组织或者用 FastPrep 仪器每次 100 mg 批量匀浆（见方案 2）。

大多数组织的处理采用上述任何一种方法均可。关于这些仪器的信息和常用的组织捣碎方法，见**动物细胞和组织的匀浆**信息栏。

d. 对于坚硬的组织或者超过 500 mg 的组织，使用 TRIzol 溶液匀浆之前，用一个预冷（使用液氮）的研钵和研杵磨碎组织。

e. 将 TRIzol 转移到一个 15 ml 带旋盖的聚丙烯离心管中。

对于悬浮培养的细胞：

a. 在预冷的离心机中，4℃下 400 *g* 离心 10 min，收集细胞。

b. 通过吸取除去培养基，将细胞团块轻轻的重悬于 0.5 倍原始培养物体积的冰冷的 PBS。

c. 离心使细胞重新成团，除去 PBS，每 5×10^6 个细胞中加入 1 ml TRIzol 溶液。

d. 迅速将细胞悬液转移到一个 15 ml 带旋盖的聚丙烯离心管中，室温下剧烈

涡旋振荡 1 min。

对于单层培养细胞：

a. 除去培养基并且用冰冻的 PBS 洗涤单层细胞 2 次。

114

b. 迅速除去 PBS，并向每 10 cm^2 单层培养物中加入 1 ml TRIzol 溶液。

c. 使用塑料刮铲将细胞刮到 TRIzol 溶液中。

d. 迅速将细胞裂解物转移到一个 15 ml 带旋盖的聚丙烯离心管中，室温下剧烈涡旋振荡 1 min。

不要节省 TRIzol 溶液。建议的用量是克服残留的 PBS 和培养基的缓冲影响所必需的。一些类型的哺乳动物细胞含有特别坚硬的结构（如黑素体），既不能在细胞裂解时被破坏，也不能在层析时去除。进行步骤 2 之前，将裂解物在 65℃下孵育 5 min可以去除这些复合物。

2. 室温下静置离心管 2～3 min，然后加入 0.2 倍体积的氯仿。不能使用 200 μL 的吸头，要用 1 ml 的带滤芯吸头，防止氯仿“爬进”过滤器。倒转密封的离心管剧烈摇动 15～20 s。

3. 室温下静置离心管 5 min，4℃下 12 000 *g* 离心 15 min。

4. 用 Pasteur 吸管将上清液转移到一支新的聚丙烯管中，记录下上清液的体积。

为减少 DNA 污染的可能性，要小心避免有机相（下层）和水相（上层）之间的白色界面。如果这个步骤很难做到，考虑使用 Phase Lock Gel 进行有机相和水相的分离（见方案 9 中的替代方案或者 www. brinkman. com）。

5. 涡旋振荡时，向上清液中缓慢加入 1 倍体积的 70%乙醇。

逐滴加入乙醇，并确保下一滴加入之前乙醇和上清液充分混合。如果混合不充分，在局部高浓度乙醇中 RNA 将沉淀出来。加入结束后，混合物中的最终乙醇含量应为 35%。

6. 将样品加入 RNeasy 离心柱中。预期的总 RNA（按照生产商的推荐方案）决定选择的离心柱。RNeasy Midi 离心柱的最大结合容量是 1 mg（相当于 5×10^6～1×10^8 个细胞或者 20～250 mg 组织），而 RNeasy Maxi 离心柱则是 6 mg（相当于 5×10^7～5×10^8 个细胞或者 0.15～1 g 组织）。

7. 室温下，将 RNeasy 离心柱在配有甩平式转头的医用台式离心机中以 3000 *g* 离心 5 min。

8. 将流出液加到同一根柱子的上部，并再次以 3000 *g* 离心 5 min。

在第一次离心中，很大一部分 RNA 不能被柱中介质俘获。

离心后，检查在离心柱的尖底部有无可见的缓冲液。如果有缓冲液存在，用吸管从收集管中除去流出液后重复离心步骤。

9. 丢弃来自柱中的流出液，对于 Midi 柱，加入 4 ml（15 ml Maxi）RW1 缓冲液到柱子的顶部。将柱子在室温下以 3000 *g* 离心 5 min。

115

10. 丢弃来自柱中的流出液，对于 Midi 柱，加入 2.5 ml（10 ml Maxi）RPE 缓冲液到柱子的顶部。将柱子在室温下以 3000 *g* 离心 5 min。

11. 重复一次步骤 10。柱子以 3000 *g* 再次离心 5 min 进行干燥。

12. 丢弃流出液，对于 Midi 柱，将其放入新的 15 ml 离心管中（Maxi 柱使用 50 ml 离心管）。

13. 为了洗脱 RNA，向 Midi 柱的顶部加入 150～200 μl H_2O（0.8～1.2 ml

Maxi)。室温下将柱子竖直放置 1 min，然后以 3000 *g* 离心 5 min。

14. 重复步骤 13 并合并洗脱液。
15. 使用 260 nm 和 280 nm 处的吸收光谱或者用 RiboGreen 染色后的荧光进行 RNA 定量（见 **RNA 定量**的信息栏）。
16. 根据在 1.5 ml 离心管中进行的各个标记反应需要的用量，准备等份量的 RNA 并另加 10%（为步骤 15～20 估计的损失）。
17. 向每份中加入 1/10 体积的 3 mol/L 醋酸钠（pH 5.2）以沉淀 RNA。短暂涡旋以混合溶液，然后向每支管中加入 2.5 倍体积的无水乙醇，再次涡旋混合。
18. 在干冰上冷冻 RNA 制备物直至冻结。
19. 融化，4℃下以 12 000 *g* 离心 15 min 回收 RNA。
20. 用 75%乙醇洗涤 RNA 团块，4℃下以 12 000 *g* 离心 5 min 回收 RNA。

    RNA 团块可以安全地储存在－80℃下 75%的乙醇中。

21. 标记前一天，4℃下以 12 000 *g* 离心 15 min 使 RNA 重新成为团块。小心除去全部乙醇，并将 RNA 在中等加热（42℃）的旋转蒸发器中干燥最多 2 min。将 RNA 重悬于 450 μl DEPC 处理的 H_2O 中。

    过度干燥可能导致 RNA 难以重新溶解，残留有乙醇也会干扰标记反应。

22. 根据生产商的说明，使用 Microcon YM-30 过滤装置将 RNA 浓缩到希望的体积（17 μl，取决于使用的标记方案）。需要时检查过滤器以确定浓缩速率。

    这一步骤除去了任何残留的盐和可能会抑制反转录反应的小分子和中等分子。

    为进行标记，总 RNA 的浓度要求为 7 mg/ml。

23. 将 RNA 用于微阵列实验前，用甲醛-1%琼脂糖凝胶进行电泳，检查 RNA 的完整性（见**哺乳动物、植物和细菌 RNA** 的信息栏）。

    从动物细胞和组织中，通过 TRIzol 单次抽提后纯化的总 RNA 制备物得到的探针与寡核苷酸阵列杂交时，有时会产生高的背景。这个问题常常通过用 TRIzol 按照下述方法再次抽提 RNA 得到解决：将 RNA 溶解在 50～100 μl H_2O 中。加入 1 ml TRIzol，然后按照生产商推荐的方法抽提 RNA。按照步骤 21 所述，再次溶解 RNA。

从数量有限的组织（<50 mg）中分离 RNA

一些 cDNA 微阵列实验，如使用氨基-烯丙基染料进行直接标记的（方案 14），要求不使用扩增而从数量有限的组织中分离大量的 RNA。

从少量组织中成功抽提 RNA 依赖于在合适的缓冲液中充分匀浆组织。对于显著纤维性而难以匀浆的组织，本方案中描述的使用 TRIzol 然后是 RNeasy 柱纯化的方法是合适的。为了避免在 TRIzol 处理时潜在的组织损失，建议使用 RLT 缓冲液进行匀浆（下面的替代方案，从步骤 4 开始）。注意要使用足够量的 RLT 缓冲液，例如，最多 50 mg 的组织可以悬浮在 600 μl 的 RLT 缓冲液中。

RNA 的产量随起始材料的数量和组织类型以及组织的转录活性而变化。通常，从每 1 mg 组织中可以获得约 1 μg 的 RNA。来自各种活组织检查形式的组织近似数量的例子如下：图 3-3 中分别示例了内窥镜活检和穿刺活检（trucut biopsy）。

活组织检查类型	组织重量
内窥镜活检（标准的穿刺钳）	约 10 mg/活检
穿刺活检［使用 bioptigun（14 号）的芯样］	约 100 mg /厘米长度
细针吸取（FNA）	<5 mg/活检

该信息栏由 Alex Boussioutas（Peter MacCallum Cancer Research Institute，Melbourne，Austrlia）提供。

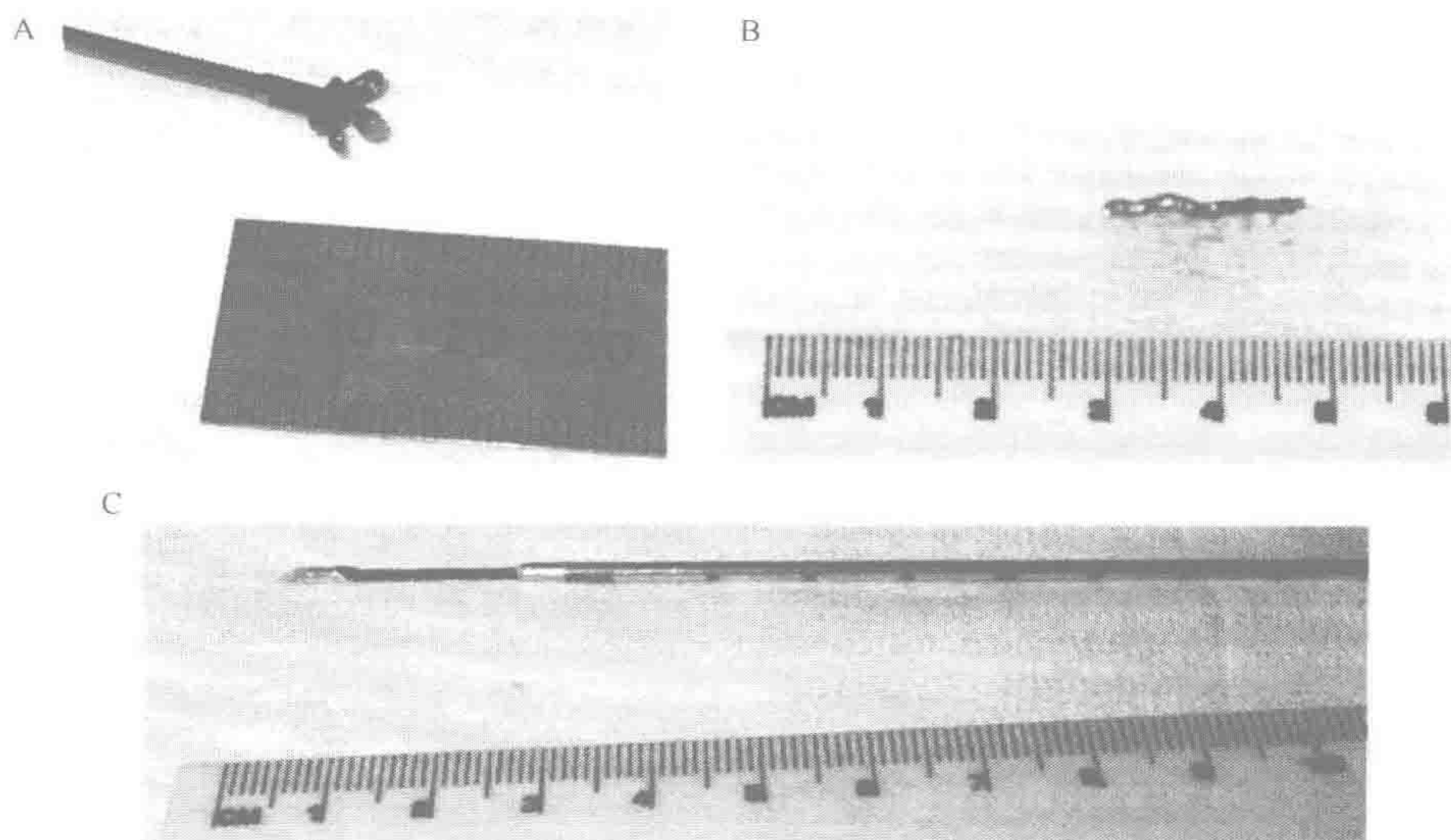


图 3-3 活组织检查样品

A. 内窥镜活检；B. 穿刺肝脏活检；C. 穿刺针

替代方案：从 OCT 包埋的病理样品中提取 RNA

Sophie Katsabanis（Peter MacCallum 癌症研究所，墨尔本，澳大利亚）

常用于包埋组织的最适切片温度（OCT）复合物主要是由糖类和聚乙烯醇组成。有许多方法能够从 OCT 包埋的组织中得到高质量的 RNA，需根据下游应用的性质选择合适的方法。

例如，对 RT-PCR 分析，从一些合并的切片能够获得足量的 RNA。Irving 等（1996）描述了一种方法，5 片 20 μm 的冰冻切片在－20℃下冻存于 GITC（异硫氰酸胍）溶液中。解冻后，样品被匀浆，通过 CsCl 超速离心分离 RNA。在 Nichols 等（1997）使用的修改过的方法中，－20℃下从 OCT 块上切取 5 张冷冻组织切片（10 μm 厚）并裱贴在玻璃载片上。室温下在空气中干燥 5 min 后，用无菌的剃须刀片将 OCT 手工刮下，然后将组织转到抽提管中。用 QIAGEN RNeasy 试剂盒从合并的切片中分离总 RNA。

使用大量 RNA 的下游应用（如 Northern 和微阵列）则需要不同的方法。可以选择的方法有多种，取决于能否获得必需的仪器。任何一种方法的主要目的都是在进行 RNA 分离前尽可能地除去 OCT 复合物。任何一种基于水的商品化试剂盒（QIAGEN RNeasy，Ambion RNAqueous）都能用来分离 RNA。GITC 破碎和 CsCl 超速离心也有同样的效果，尽管由于离心转子的容量小和离心时间的延长将会影响产量。

为了获得大量 OCT 包埋的组织，用一台冷冻的（−80℃或更低）显微镜用薄片切片机从整个组织块切取 100 μm 厚的切片直到仅剩下 OCT，将其丢弃。组织切片放入 GITC 中，根据方案 8 描述的方法用 CsCl 超速离心分离 RNA。作为替代，切片可以直接放入 RLT 缓冲液（来自 QIAGEN RNeasy 试剂盒）。残留的 OCT 将溶解在缓冲液中，因为 RLT 是一种水溶液。短暂离心使组织碎片变成团块，除去 RLT 并更换新鲜的 RLT 缓冲液。这是一个重要的步骤，因为从 OCT 释放出的大量糖类将阻塞 RNeasy 柱，导致 RNA 产量减少。RNeasy 的方案根据供应商的说明进行。

附加材料

冷冻的显微镜用薄片切片机（−80℃或更低）

含有 β-巯基乙醇的 RLT 裂解缓冲液〈!〉（RNeasy 试剂盒中提供）

方法

1. 用剃须刀片从组织中移走 OCT 块。

   建议切除的时候将包埋块保持在−80℃，以减少组织的融化和由此导致的 RNase 活性造成的 RNA 降解。如果这不能做到，取出保存在液氮中的 OCT 并稍作融化以易于切片。

2. 在 50 ml falcon 管中，将组织放入 RNA*later*。

   对于一个 10 mm×10 mm 的 OCT 块，使用 5 ml RNA*later*。

   包埋组织不超过 5 mm 厚很重要，否则 RNA*later* 溶液不能足够快地渗透去保护整个样品，冰冻组织中的扩散也会受阻。为了得到最佳结果，尤其是包埋的组织块大小未知，或者已知尺寸大于 5 mm，包埋块应该纵向切开。这样就暴露了又一个组织截面，使得 RNA*later* 能较快地扩散到组织中。

3. 剧烈涡旋振荡在 RNA*later* 溶液中的组织块，直到 OCT 开始溶解。

   OCT 不会完全溶解，相反，一种类似于外壳或皮肤的物质将残留在组织周围。OCT 需要足够地软，以便于用镊子/钳子从其中将组织物理分离。

4. 移开组织并放在新的试管中，称重样品。
5. 加入适量含有 β-巯基乙醇的 RLT 裂解缓冲液（参考 RNeasy 的方案），然后用 Ultraturax 或类似的转子—定子匀浆器进行组织匀浆。
6. 将匀浆转移入一个 15 ml 带旋盖的离心管中。
7. 继续用 RNeasy 柱纯化 RNA，从主要方案的步骤 2 开始。

方案 2

118

用 FastPrep 仪器从组织中分离总 RNA

FastPrep FP 120 仪器（Q-Biogene 6000-120）用来匀浆质量少于 100 mg 的组织。在一个耐撞击的含有裂解介质且用于机械性破碎和匀浆组织的 2 ml BioPulverizer 管中破碎组织样品，并在 1 ml TRIzol 中匀浆。这个仪器最多可以摇动 12 支试管，在非常高的速度下进行转动和投掷运动。匀浆通常在数秒钟内完成，但坚韧的样品如小鼠尾巴或骨骼可能需要较长的匀浆时间。由于每种样品放在各自单独的试管中，避免了样品之间的交叉污染。

材料

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

单相裂解溶液（TRIzol 15596，GIBCO）

见方案 1 中表 3-3 和表 3-4。

专用设备

含有裂解介质 D 的 BioPulverizer 管（2 ml）（6513-001，Q-Biogene Inc.，Carlsbad California）

FastPrep FP 120 仪器（6000-120，Q-Biogene Inc.，Carlsbad California）

附加试剂

本方案可能需要方案 1 中列举的试剂。

方法

1. 将 FastPrep FP 仪器预置在 6 m/s，时间为 25 s。
2. 测量组织的质量。在一个 2 ml BioPulverizer 管中不要使用多于 100 mg 组织。
3. 向每支 BioPulverizer 管中加入 0.5 ml TRIzol。将试管在冰上存放 5 min。
4. 迅速将新鲜切下的组织转移到含有 TRIzol 的 BioPulverizer 管中，确保管盖旋紧。

   需要约 0.5 ml 的空气。较少的空气会导致由于试管损坏或管盖周围变形引起的样品泄漏。

5. 将试管放入 FastPrep 仪器中，确保管盖锁的尖头直接位于样品管上方，向下旋紧尖头。

119

6. 关上盖子开始操作。
7. 操作完成后，立即移走试管并将其在冰上放置 2 min。
8. 如果组织没有完全匀浆，重复匀浆步骤 2 或 3 次。每次摇动后要将试管在冰上放置 2 min，否则因为 TRIzol 溶液的热膨胀试管可能会爆裂。
9. 当匀浆完成后，将试管从 FastPrep 仪器中转移到冰桶中。立即向匀浆中加入 0.5 ml TRIzol，充分混合溶液。
10. 随后进行方案 1 中的步骤 2～18。

120

方案 3

用 TRIzol 从植物组织中分离总 RNA

Damares Monte[*†]，**Shauna Somerville**[*]［[*]华盛顿卡内基研究院，斯坦福，加利福尼亚，94305；[†]E. M. B. R. A. P. A.（巴西农业研究公司）遗传资源与生物技术研究所，巴西利亚，巴西］

你为了跟上地球和她的灵魂而工作着。因为虚度将会成为机遇的陌生人，并且偏离生命的队伍，这支队伍向着神圣前进，并且自豪地顺从于无限。

Kahlil Gibran

高通量的 DNA 微阵列技术（Schena et al. 1995）已经生成了大量生长在不同条件下的多种植物的基因表达谱的目录（关于综述，请见 Kehoe et al. 1999；Richmond and Somerville 2000）。最早的阵列通常仅仅含有几百个基因（Ruan et al. 1998；Aharoni et al. 2000；McGonigle et al. 2000；Reymond et al. 2000）。但是，随着植物基因组测序的不断进展，点样在阵列上的基因已经扩展到几千个（Arabidopsis Functional Genomics Consortium，http://afgc.stanford.edu）（Harmer et al. 2000；Maleck et al. 2000；Wang et al. 2000；Zhu and Wang 2000）。

从植物组织制备微阵列分析要求质量的 RNA 要比从动物组织更加困难。在抽提和纯化时必须非常认真，因为植物 RNA 可能被多酚、多糖、蛋白质或 DNA 污染，所有这些都将抑制体外 cDNA 的合成和标记。这些污染物质和（或）小的核酸片段（＜200 bp）的存在可能导致杂交背景明显增加。许多植物 RNA 转录物中固有的二级结构单元可能会阻止反转录酶的接近或者导致 cDNA 合成的过早终止（Zhu and Wang 2000）。这些问题，单独或同时存在，可能会使本来已经有限的微阵列实验的动态范围（Bertucci et al. 1999；Lipshultz et al. 1999）减小至低丰度的转录物无法检测到。

由于多酚和多糖结合到 RNA 上并且危及微阵列分析的整体性能，因此选择从特定的植物物种或组织中分离和纯化 RNA 的方法是成功的一个关键步骤。本方案和方案 4 中描述了两种纯化总 RNA 方法，用于获得高质量的植物 RNA。方案 3 描述了如何使用 TRIzol 方法从多种软的植物组织中纯化 RNA；方案 4 描述 Chang 等（1993）提出的松树方法，推荐用于从富含多酚和蛋白聚糖的顽拗植物（recalcitrant plant）物种和木质化组织中抽提 RNA。尽管起源于从松果中分离 RNA，本方法仍可成功地用于从一些植物物种如拟南芥（*Arabidopsis thaliana*）中分离高纯度的 RNA。根据植物组织的来

121

源和类型，两个方案的 RNA 产量变化较大。本方案通常从 1 g 拟南芥叶组织中可以得到 1.7 mg 的总 RNA，而方案 4 可从松树种子和松果中得到约 240 μg 的 RNA。对于小规模制备，使用商业化的试剂盒也是一种替代方法（如来自 QIAGEN 的 RNeasy Plant Kit 试剂盒）。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

重要事项：本方案中所有试剂使用经 DEPC 处理的 H_2O 配制（见 **RNA 操作的特别注意事项**信息栏）。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

氯仿〈!〉

焦碳酸二乙酯（DEPC）〈!〉

乙醇（95%和 75%）

据报道，某些批次的乙醇（100%或 200 proof）中含有会使微阵列试验中产生很深荧光背景的污染物质，95%的乙醇不存在这个问题。为避免这个问题，可以用 95%的乙醇或者已证明不会产生高荧光背景的无水乙醇。Warner-Graham 公司（Cockeysville，Maryland）出售的无水乙醇（USP 级乙醇，64-17-5）效果很好。

异丙醇（2-丙醇）

液氮〈!〉

单相裂解溶液（如 TRIzol；见方案 1 中的表 3-3 和表 3-4）

优先选择的 RNase 抑制剂（见**抑制 RNase** 的信息栏）

RNA 沉淀溶液

0.8 mol/L 柠檬酸钠

1.2 mol/L NaCl

不需要进行 pH 调节

离心机和转子

医用离心机（Eppendorf 5810R 或相当的）

Eppendorf 转子（F-34-6-38 或相当的）

专用设备

研钵和研杵（在液氮中预冷）（Fisher Scientific）

Polytron 匀浆器

Polytron 用来破碎细胞和多种类型的组织（Kinematica Model 1 PT3100，配备 Probe 9100071；Brinkmann）。使用前，清洁 Polyton 使其没有 RNase。拆开探针，先用水洗涤部件，然后用无水乙醇，最后用可降解 RNA 的 RNaseZap 处理。组装探针并在水中匀浆。

水浴，预设为60℃

122

方法

重要事项：在开始本方案的操作之前将离心管存放在液氮中。此外，如果同时处理多个样品，在处理其他样品时将装有粉状组织的离心管保存在液氮中。

1. 按照表3-1中描述收获和储存植物组织。在液氮存在的条件下，用预冷过的研钵和研杵把1 g组织研成粉末。
2. 将粉末状组织转移到一支在液氮中预冷的50 ml无菌带旋盖塑料离心管中。在处理其他样品时将离心管保存在液氮中。
3. 从液氮中取出含有粉状植物组织的离心管，立即加入15 ml TRIzol试剂。
4. 剧烈倒转离心管，使粉末状植物组织和TRIzol裂解溶液充分混合，确保当粉末状组织刚融化时就和试剂接触。

   如果不能被TRIzol溶液中的RNase抑制剂中和，结合在膜相中的RNase活性会非常高。

   警告：非常小心地操作含有缓冲液和粉末状组织的聚乙烯管。溶解氮的释放可能会冲掉管盖，因而在操作匀浆缓冲液的所有阶段要戴上安全眼镜。

5. 如果组织是坚硬或纤维性的，将裂解物在室温下孵育5 min。
6. 用Polytron匀浆器匀浆组织2次，每次15 s。在摇动和涡旋的间隔给试管放气一次。
7. 将样品在4℃下以10 000 *g*（9000 r/min）离心10 min。
8. 将上清液转移到一支新鲜灭菌的50 ml无菌带旋盖塑料离心管中，弃去团块。
9. 向离心管中加入3 ml氯仿。
10. 剧烈摇动试管并涡旋15 s使两相混合。在摇动和涡旋的间隔给试管放气一次。
11. 将试管在室温下放置2～3 min。
12. 4℃下以10 000 *g*（9000 r/min）离心15 min进行相分离。
13. 小心地将水相转移到新的50 ml无菌带旋盖塑料离心管中。注意不要将界面处和下层相中的物质带入。弃去界面相和下层有机相。

    如果本步骤操作困难，可以考虑使用锁相胶进行有机相和水相的分离（见方案9中**替代方案：使用锁相胶纯化RNA**的信息栏）。

14. 测量水相的体积，然后加入0.5倍体积的异丙醇和0.5倍体积的RNA沉淀溶液以沉淀RNA。盖上管子轻微反转混合。
15. 将试管在室温下放置10 min。

123

16. 4℃下以10 000 *g*（9000 r/min）离心10 min，回收RNA。
17. 小心倾去上清液，然后用20 ml 75%的乙醇洗涤RNA团块。短暂涡旋清洗管壁。
18. 4℃下以10 000 *g*（9000 r/min）离心10 min，回收RNA。弃去上清液，将开盖的试管在化学通风橱中竖立5～10 min（时间不要长!）直到没有一丝可见的上清液。

19. 用吸管上下吹打数次，将 RNA 重悬于 250 μl DEPC 处理的水中。
20. 室温下将样品转移到微量离心管中，在台式微量离心机中以最高速离心 5 min，除去不能溶解的物质。
21. 将 RNA 溶液（上清液）转移到一个新的试管中。通过在 260 nm 和 280 nm 的吸收光谱或使用 RiboGreen 染色后的荧光对 RNA 进行定量（见 **RNA 的定量**信息栏）。
22. 选用 **RNA 的保存和回收**信息栏中的一种方法储存 RNA。（可选择的）加入一种 RNase 抑制剂（如 RNasin），根据生产商者的说明（Sambrook and Russell 2001）进行操作。

    将 RNA 用于微阵列实验前，通过甲醛-1%琼脂糖凝胶电泳检查 RNA 的完整性（见**哺乳动物、植物和细菌 RNA** 的信息栏）。

方案 4

124

松树方法分离植物 RNA

Damares Monte[*,†]，**Shauna Somerville**[*] [[*]华盛顿卡内基研究院，斯坦福，加利福尼亚，94305；[†]E. M. B. R. A. P. A.（巴西农业研究公司）遗传资源与生物技术研究所，巴西利亚，巴西]

多种植物的组织中含有大量可能会污染 RNA 制备物的多酚复合物、脂多糖和蛋白质。这会引起乙醇沉淀后 RNA 溶解的困难，减弱了 cDNA 合成和标记的效率，并产生杂交背景。

根据 Schneiderbauer 等（1991）和 Katterman 与 Shattuck（1983）从富含核糖核酸酶、多糖和多酚的松树幼苗中分离总 RNA 的原始方法，Chang 等（1993）提出了松树方法。利用传统的苯酚方法从这类型的植物组织中抽提 RNA 的产量通常较低，并存在氧化的酚复合物和 RNA 牢固结合的浅褐色沉淀。在裂解缓冲液中加入与污染物结合的聚乙烯吡咯烷酮（PVP）、β-巯基乙醇（一种还原剂）和一种阳离子去污剂溴化十六烷三甲基铵（CTAB）可以克服这个问题。使用氯仿代替苯酚除去蛋白质。

本方法避免了使用苯酚、盐酸胍或者异硫氰酸胍，并且不要超速离心。回收的 RNA 通常是完整和无色的，每克组织产量为 100～240 μg，并且 A_{260}/A_{280} 比值在 1.7 以上。因为本方法需要用氯化锂进行长时间的 RNA 沉淀操作，根据制备的 RNA 的量，完成这个方案的时间从 4 h 至过夜。暴露在氯化锂中的时间越长，回收的 RNA 的量越多。如果经过测试没有 DNA 污染，RNA 可以直接用于 DNA 微阵列杂交和作为纯化的 poly (A)$^+$ RNA 的来源用于制备标记的靶。

材料

125

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

重要事项：本方案中所有试剂均使用经 DEPC 处理的 H_2O 配制（见 **RNA 操作的特别注意事项**信

息栏）。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录2。

使用前将储存液稀释到合适的浓度。

氯仿：异戊醇（24：1）〈!〉

抽提缓冲液

2%（*m/V*）CTAB

2%（*m/V*）PVP K 30

100 mmol/L Tris-HCl（pH 8.0）

25 mmol/L EDTA

2.0 mol/L NaCl

0.5 g/L 亚精胺

2% β-巯基乙醇〈!〉

混合除β-巯基乙醇外所有的成分。将溶液分成小份，采用高压蒸汽灭菌15 min。使用前向抽提缓冲液中加入β-巯基乙醇。

LiCl（8 mol/L，相对分子质量为42.39）

将33.9 g氯化锂加入到50 ml水中，并将溶液体积调节到100 ml。氯化锂全部溶解后，采用液体循环程序高压灭菌15 min。

SSTE

1 mol/L NaCl

0.5%（*w/V*）SDS〈!〉

10 mmol/L Tris-HCl（pH 8.0）

1 mmol/L EDTA（pH 8.0）

离心机和转子

医用离心机（Eppendorf 5810R或相当的）

Eppendorf转子（F-34-6-38或相当的）

专用设备

研钵和研杵（在液氮中预冷）（12-061-5B，Fisher Scientific）

水浴，预设为65℃

方法

1. 将15 ml抽提缓冲液（含有β-巯基乙醇）加到一支50 ml的聚丙烯管中。密封试管并在水浴中将缓冲液加热到65℃。
2. 在液氮存在下，用研钵和研杵把2～3 g植物材料研成细粉。
3. 迅速将组织粉末加入到15 ml预热的抽提缓冲液中，倒转试管数次将粉末和缓

冲液完全混合。

警告：非常小心地操作含有缓冲液和粉末状组织的聚乙烯管。溶解氮的释放可能会冲掉管盖，因而在操作匀浆缓冲液的所有阶段要戴上安全眼镜。

4. 用等体积的氯仿：异戊醇（24：1）抽提组织溶液2次，室温下以1000 *g*（9000 r/min）离心5 min进行相分离。回收上（水）层时，一定要避免带入界面。第二次抽提后，将水相转移到一支新的50 ml离心管中。

   如果回收不含界面物质的水相很困难，可以考虑使用锁相胶进行有机相和水相的分离（见方案9中**替代方案：使用锁相胶纯化RNA**的信息栏）。

5. 向水相加入0.3倍体积的8 mol/L LiCl。倒转试管数次以混合溶液。将试管在4 ℃下静置12～16 h。4℃下以10 000 *g*（9000 r/min）离心20 min收集RNA。

   LiCl是强的脱水剂，它降低RNA的溶解性（Hearst and Vinograd 1961 a，b）并且从染色质上除去蛋白质（Kondon et al. 1979）。高分子质量的RNA能被高浓度（≥0.8 mol/L）的LiCl沉淀并通过低速离心回收（Barlow et al. 1963）。经过1 h、2 h和6 h沉淀后，RNA的产量分别为过夜沉淀产量的30％、65％和90％。

6. 将RNA团块溶解于500 μl SSTE中。

   如果要进行poly（A）$^+$ RNA的抽提，将RNA溶解在0.5％ SDS中替代SSTE，然后前进到方案7。否则，前进到本方案的步骤7。

7. 用等体积的氯仿：异戊醇（24：1）抽提RNA溶液一次。4℃下以1000 *g*（9000 r/min）离心10 min进行相分离。将上层（水）相转移到一支新的50 ml离心管中。

8. 加入2倍体积的乙醇，在－80℃沉淀RNA至少30 min，或者在－20℃下2 h。

9. 在微量离心机中以最大速度离心20 min收集RNA。吸去上清液并将打开盖子的离心管放在化学通风橱中数分钟，直到最后一丝可见的上清液蒸发掉。将RNA团块重悬在经过DEPC处理的水中。预期的总RNA产量为每克原始材料100～240 μg/g。以约7 mg/ml的浓度溶解RNA。

10. 将RNA溶液转移到一支新的试管中。通过在260 nm和280 nm的吸收光谱或使用RiboGreen染色后的荧光对RNA进行定量（见**RNA的定量**信息栏）。

11. 选用**RNA的保存和回收**信息栏中的一种方法储存RNA。（可选择的）加入一种RNase抑制剂（如RNasin），根据生产商的说明（Sambrook and Russell 2001）进行操作。

    将RNA用于微阵列实验前，通过甲醛-1％琼脂糖凝胶电泳检查RNA的完整性（见**哺乳动物、植物和细菌RNA**的信息栏）。

方案5

用FastTrack分离poly（A）$^+$RNA

Jun S. Wei，**Javed Khan**（先进技术中心，儿科肿瘤学部门，国家癌症研究所，国立卫生院，盖瑟斯堡，马里兰州，20877）

在基因表达谱分析实验中，使用总 RNA 而非 poly (A)$^+$ RNA 作为产生探针的模板时获得较强的信号和较低的背景（Wildsmith et al. 2001）。较差的信噪比常常是由于通过 oligo (dT)-纤维素纯化 mRNA 时的固有的丢失。当细胞和组织的数量有限时，这是一个非常关键的问题。但是，使用 poly (A)$^+$ RNA 有两个潜在的优势：首先，用总 RNA 作为模板时，标记反应的特异性完全依赖于 oligo (dT) 选择 mRNA 作为模板的能力。因为荧光标记 cDNA 的合成很少能够延伸到 mRNA 的全长，当在微阵列上的杂交的靶对应于转录物的 3′端时，用 oligo (dT) 作为引物是合适的。如果可以得到纯化的 poly (A)$^+$ RNA，其他标记选择也成为可能。例如，用寡核苷酸随机引物作为引物合成 cDNA 的第一链，它可以在 mRNA 的全长上引发合成并且产生的 cDNA 片段虽然较小，但是代表了大部分的 mRNA 序列。第二，用来分离 poly (A)$^+$ RNA 的方法——用 oligo (dT) 小珠装柱的亲和层析——提供了另一方面的纯化，能用来去除用单相裂解试剂抽提组织时随 RNA 同时分离的污染物如多糖和多酚。

本指南中该方案和方案 6 供希望用 poly (A)$^+$ RNA 作为合成荧光标记靶标的模板的研究人员参考。本方案中使用商品化的试剂盒（FastTrack 2.0，Invitrogen 销售），而方案 6 描述了常用 oligo (dT) 层析的另一种形式。

FastTrack 2.0 可用来从细胞和组织中直接分离 poly (A)$^+$ RNA，或根据方案 1 所述制备的总 RNA。本实验是设计从 $1\times10^7\sim5\times10^8$ 个培养的哺乳动物细胞，400 mg～1 g 组织，或 0.2～1.0 mg 总 RNA 中纯化 mRNA。

材料

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

乙醇（无水）

据报道，有些批次的乙醇（100%或 200 标准强度）中含有会使微阵列试验中产生很深荧光背景的污染物质，95%的乙醇不存在这个问题。为避免这个问题，可以用 95%的乙醇或者已证明不会产生高荧光背景的无水乙醇。Warner-Graham 公司（Cockeysville，Maryland）出售的无水乙醇（USP 级乙醇，64-17-5）效果很好。

NaCl (5 mol/L)

磷酸盐缓冲液（PBS），不含 Ca^{2+} 和 Mg^{2+}（14190-136，GIBCO）

RNA*later* 稳定剂（221462，QIAGEN）

RNase*Zap*（9780，Ambion）

醋酸钠（3 mol/L，pH 5.2）

核酸和寡核苷酸

总 RNA 按照方案 1 制备（见步骤 3）

细胞和组织

哺乳动物细胞的单层或悬浮培养物或新鲜/冰冻组织（见步骤 1）

专用设备

FastTrack 2.0 试剂盒（K1593-02，K 1593-03，Invitrogen）

匀浆器

配备 Probe 9100071（Brinkmann 2713 700-8）的 Polytron PT1200 匀浆器（Kinematica AG，Switzerland）或相当的仪器，用来捣碎细胞和组织。使用前，清洁 Polyton 使其没有 RNase。拆开探针，先用水洗涤部件，然后用无水乙醇，最后用可降解 RNA 的 RNaseZap 处理。组装探针并在水中匀浆。可以使用其他相当的仪器（FastPrep 仪器）。关于各种仪器的讨论，见**动物细胞和组织的匀浆**信息栏。

研钵和研杵

捣碎大量的（>500 mg）软组织和任何数量坚硬的纤维性组织时所必需的（见步骤 1）（12-061-5B，Fisher Scientific）。

方法

1. 准备分离 RNA 的细胞或组织样品。

对于组织：组织匀浆前尽可能将组织保持在－80℃冷冻或在液氮、RNA*later* 中直到匀浆。

a. 组织称重。

b. 将 100 mg 组织块直接放入装有 FastTrack 2.0 裂解缓冲液的 50 ml 圆锥形聚丙烯试管中。每毫升裂解液适用于大约 60 mg 组织。

c. 立即用带有旋转刀片的组织匀浆器（Ploytron）匀浆组织 30 s。

少于 500 mg 的软组织容易匀浆，大量的组织可以按 500 mg 分批处理。硬的组织要用液氮预冷的研钵和研杵研磨成粉末，然后将粉末用合适体积的 FastTrack 2.0 裂解缓冲液在 Polytron 匀浆器中匀浆 30 s。

d. 匀浆后，在一个含有 2 ml 裂解缓冲液的 15 ml 锥形试管中操作 Polytron 探针，并将溶液加入到组织裂解物中。如果仍可以见到颗粒状的物质，4℃下以 12 000 *g*（9000 r/min）离心样品 30 min。

悬浮培养的细胞：

a. 4℃下在预冷的离心机中以 400 *g* 离心 10 min，收获细胞。

b. 吸走培养基，将细胞团块重悬于 0.5 倍原始培养物体积的冰冷的 PBS。

c. 与步骤 a 中同样地离心收获细胞，除去 PBS，并按每 10^6 个细胞加入 1 ml 的 FastTrack 2.0 裂解缓冲液。

d. 立即将细胞悬液转移到聚丙烯试管中，并将细胞裂解物反复通过 21 号针数次以破碎细胞和剪切基因组 DNA。

本步骤包括将裂解物通过针吸入注射器并从注射器再通过针推出，重复该过程 5 次。作为替代，室温下将细胞在 Polytron 匀浆器中匀浆 15～30 s。匀浆后，在一个含有 2 ml 裂解缓冲液的 15 ml 锥形试管中操作 Polytron 探针，并将溶液加入到组织裂解物中。

单层生长的细胞：

a. 用冰冷的 PBS 冲洗细胞两次，并向每 10^6 个细胞加入 1 ml FastTrack 2.0 裂解缓冲液。

b. 立即将细胞刮到缓冲液中，并将细胞悬液转移到聚丙烯试管中。将细胞裂解物反复通过 21 号针数次以破碎细胞和剪切基因组 DNA。

本步骤包括将裂解物通过针吸入注射器并从注射器再通过针推出，重复该过程 5 次。作为替代，室温下将细胞在 Polytron 匀浆器中匀浆 15～30 s。匀浆后，在一个含有 2 ml 裂解缓冲液的 15 ml 锥形试管中操作 Polytron 探针，并将溶液加入到组织裂解物中。

2. 将细胞裂解物在 45℃下孵育 30～60 min。

3. 向每 15 ml FastTrack 2.0 裂解缓冲液中加入 950 μl 5mol/L NaCl (NaCl 最终浓度为 0.5 mol/L)。如果裂解物明显地黏稠，如同本方案步骤 1 那样将裂解物通过 21 号针来剪切 DNA。前进到步骤 4。

在本步骤结束时，确保没有颗粒状物质的存在。方案下面的部分中需要加入 oligo (dT) 树脂并通过离心回收结合的 RNA。颗粒状物质将被回收到团块中并且可能危害方案中剩余的操作。

从总 RNA 分离 mRNA：最好按照每批 1 mg 来分批处理总 RNA。每批预期的 poly (A)$^+$ RNA 产量为 10～50 μg，取决于组织或细胞的来源。

a. 按照方案 1 的步骤 17～20 用乙醇沉淀总 RNA。

b. 将团块重悬在 100 μl 洗脱缓冲液中（包括在 FastTrack 2.0 试剂盒内）。

c. 将 RNA 溶液加入到含有 10 ml FastTrack 2.0 裂解缓冲液的 50 ml 离心管中。

d. 将溶液在 65℃下孵育 5 min，然后立即将离心管在冰上放置正好 1 min。

e. 将离心管在室温下放置，并加入 650 μl 5mol/L NaCl，轻微倒转混合溶液。

f. 前进到步骤 4。

4. 加入 oligo (dT) —纤维素（包括在 FastTrack 2.0 试剂盒内）。以 10^6 个培养的哺乳动物细胞或 20 mg 组织中含有 0.75 μg mRNA 来估计需要加入树脂的量，每微克 mRNA 需要 1 mg 树脂。

5. 将试管在室温下放置 1～2 min 使得树脂重新水合。涡旋样品使树脂完全重悬。将试管放到一个摇动的装置上并且在室温下轻轻地水平摇动 60～90 min。

6. 室温下以 3000 *g* 离心浆状树脂 5 min。在本步骤和以后全部的离心中，要确保闸是关闭或者设定为“低”。轻轻地吸走上清液，尽量不要影响到团块。

7. 将 oligo (dT) 树脂轻轻地重悬在 20 ml 结合缓冲液中（包括在 FastTrack 2.0 试剂盒内）。重复步骤 6。

8. 将 oligo (dT) 树脂轻轻地重悬在 10 ml 结合缓冲液中。重复步骤 6。

9. 将 oligo (dT) 树脂轻轻地重悬在 10 ml 低盐清洗缓冲液中（包括在 FastTrack 2.0 试剂盒内）。重复步骤 6。

10. 重复步骤 9 直到离心后在缓冲液中没有气泡存在。步骤 9 可能需要重复 5 次。

11. 用一个血清移液管将最少数量的树脂重悬在低盐缓冲液中。将树脂转移到放置在微量离心管内的离心柱中（包括在 FastTrack 2.0 试剂盒内）。以 5000 *g* 将

微量离心管内的离心柱离心 10 s，弃去离心管中的液体。

12. 重复步骤 11 直到全部的树脂被转移到离心柱中。
13. 通过以吸头混合树脂的方式，用大约 500 μl 低盐清洗缓冲液来清洗离心。小心不要损坏膜！在 5000 *g* 将离心柱离心 10 s。读取离心管中液体的 OD_{260} 值。 131
14. 重复步骤 13 直到流出液体的 OD_{260} 值小于 0.05，然后将柱子放入到一个新的微量离心管中。
15. 向柱子中加入 200 μl 洗脱缓冲液（包括在 FastTrack 2.0 试剂盒内）。用吸头混合缓冲液和树脂，5000 *g* 离心 10 s。
16. 重复步骤 15 一次。微量离心管中洗出液的体积约为 400 μl。
17. 从微量离心管中回收洗出液，并加 0.15 倍体积（约 60 μl）的 2 mol/L 醋酸钠（包括在试剂盒内）和 2.5 倍体积的乙醇。
18. 将试管在干冰或－80℃冷冻直到洗出液冻结。
19. 室温下溶解洗出液，并在微量离心机中 4℃下以最大速度（16 000 *g*）离心 15 min。
20. 小心吸走全部乙醇，注意不要影响 mRNA 团块。将开口的试管竖放在化学通风橱内直至最后一丝可见的乙醇消失。
21. 将 mRNA 团块溶解在 20～50 μl 10 mmol/L Tris-Cl（pH 7.8），0.1 mmol/L EDTA 中。
22. 通过在 260 nm 和 280 nm 的吸收光谱或使用 RiboGreen 染色后的荧光对 RNA 进行定量（见 **RNA 的定量**信息栏）。
23. 选用 **RNA 的保存和回收**信息栏中的一种方法储存 RNA。

    将 RNA 用于微阵列实验前，通过甲醛-1%琼脂糖凝胶电泳检查 RNA 的完整性（见**哺乳动物、植物和细菌 RNA** 的信息栏）。

方案 6

132

分离膜结合的多核糖体 RNA

Maximilian Diehn（生物化学系，斯坦福大学医学院，斯坦福，加利福尼亚，94305）

DNA 微阵列最常用于对来自多组配对的哺乳动物细胞或组织的 RNA 抽提物中的转录物进行比较和分类。实验设计中的 一个固有假设是在总 RNA 制备物中 mRNA 的丰度可用作基因表达的准确度量。然而，由于不同的 mRNA 翻译效率各不相同，总 RNA 中转录物的丰度可能并不能反映细胞中的蛋白质合成谱。原理上，通过 DNA 微阵列来分析从多核糖体分离的 mRNA，也就是说，捕获正在翻译的 mRNA，能够得到更为精确的基因表达谱图。mRNA 另一个令人感兴趣的特征是它们的亚细胞定位。这一特性与基因表达水平和翻译速率相结合，为所编码基因产物的生物学特性提供了更深入的了解。

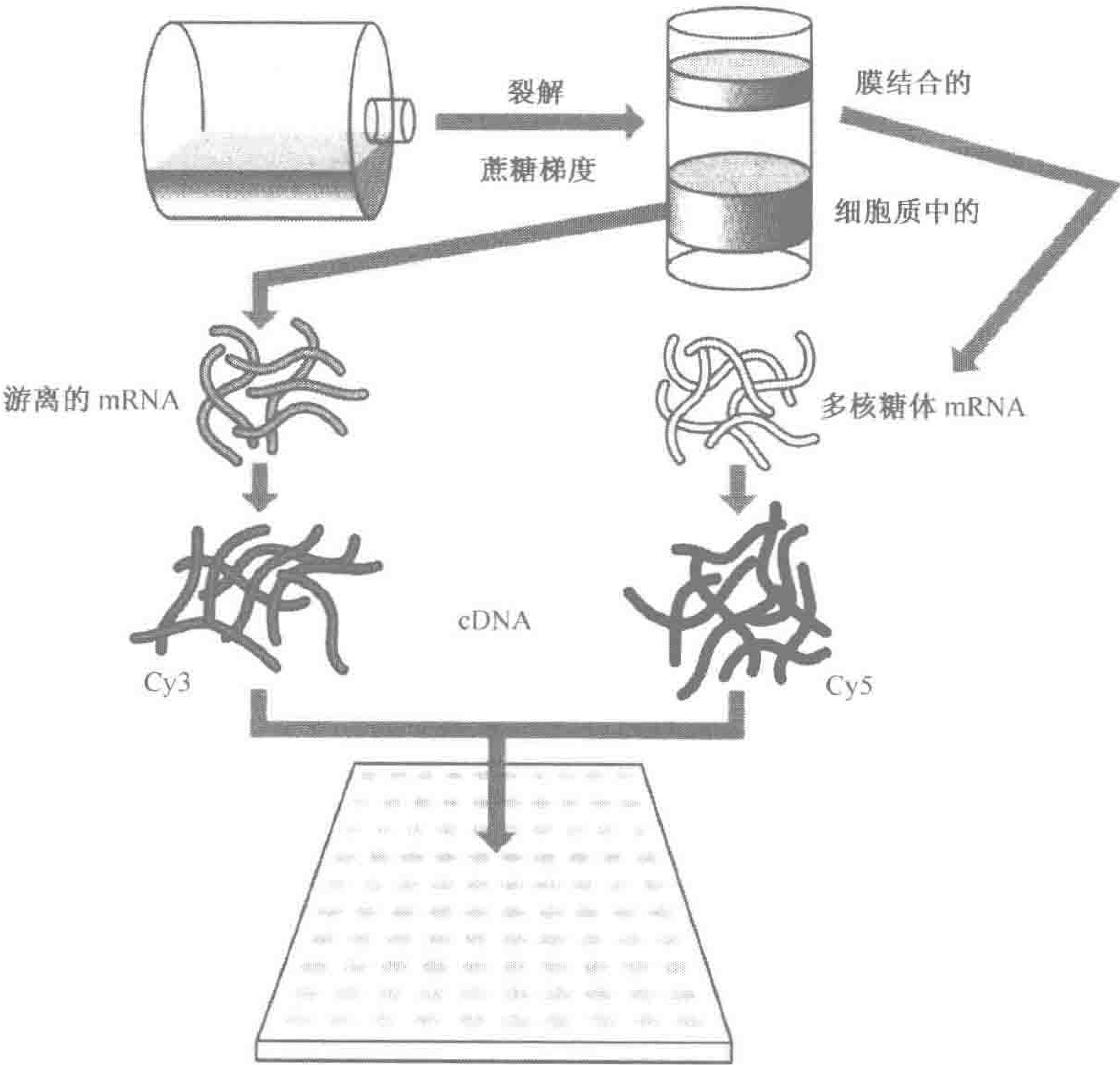
大量文献表明，在从酵母到人类的真核生物中，mRNA 在细胞中的定位对细胞或

机体的功能有着重要的生物学意义。对基因表达的全面理解需要综合关于 mRNA 的表达水平、翻译速率和它们的亚细胞定位的知识。

关于 mRNA 不同的亚细胞定位，最基本的例子就是编码细胞溶质或细胞核蛋白质的转录物与那些编码膜结合的或分泌蛋白质的转录物在空间上的分隔。属于后者的蛋白质在翻译的同时通过附着在粗面内质网（ER）的细胞质面的核糖体插入到粗面内质网中，而前者由游离在细胞溶质中的核糖体翻译。蛋白质合成在空间上的分隔形成了截然不同的 mRNA 的聚集，根据和细胞膜结构的结合与否可以将它们分开。一旦分开后，这两部分就可以和 DNA 微阵列杂交并且后续的分析能产生两种类型的重要信息：

- mRNA 编码的蛋白质的亚细胞定位可以根据它们的转录物的亚细胞定位进行推断。
- 未预料到的 mRNA 与细胞膜结构的结合可以得到证实，并且做上标记用于将来研究。

现有许多方法可以用于分离膜结合的和游离的 mRNA。一种方法是利用细胞脂类结构具有相对较低的密度通过沉降平衡（sedimentation equilibrium）把膜和密度较大的组分分离开来（Mechler 1987）。第二类方法通过使用去垢剂的差异性抽提与沉降速度离心相结合来分离两类 mRNA（Stoltenburg et al. 1995）。



133

图 3-4　从细胞系中分离膜结合多核糖体的步骤

细胞通过低渗裂解，然后通过在蔗糖梯度中的平衡密度离心将膜结合 RNA 与游离的 RNA 分离。总 RNA 的分离与含有膜结合的或游离核糖体的那部分是分开的。使用从膜结合或游离核糖体中分离出的 mRNA 来合成 cDNA，并且分别标记上荧光染料 Cy3 和 Cy5，用于和 DNA 微阵列杂交（修改过；来自 Diehn et al. 2000，经许可）。

也可以使用抗信号肽识别系统成分的抗体进行纯化的方案来分离和粗面内质网结合的 mRNA。下面的方案使用在蔗糖梯度中的沉降平衡离心。尽管实验细节是特别针对本方法的，但是在原理上，这里描述的分析方法适用于分离 RNA 各种成分的实验方法。图 3-4 中显示了该方案的整体安排。

在设计研究 RNA 的亚细胞分隔化实验时，一定要在 DNA 微阵列上包含对照探针。对应于亚细胞定位已知的转录物的探针作为对照有两个重要的原因。首先，通过检查亚细胞定位已知的 mRNA 的荧光比率分布可以评价分离过程是否成功。第二，通过观察到的已知 mRNA 的荧光比率能够鉴别出具有相似性质的其他转录物。在进行膜结合多核糖体的实验中，DNA 微阵列应该包含尽可能多的编码已知亚细胞定位蛋白质的探针。这种类型的微阵列单元至少要有数百个。

材料

134

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

重要事项：本方案中所有试剂使用经 DEPC 处理的 H_2O 配制（见 **RNA 操作的特别注意事项**信息栏）。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

氯仿（不含任何添加剂，如异戊醇）〈!〉

环己亚胺（10 mg/ml）（Sigma）〈!〉

环己亚胺在碱性溶液中不稳定。环己亚胺通过与易位酶直接作用和阻断核糖体内肽酰基 tRNA 由 A 位点到 P 位点的移动来抑制真核生物中蛋白质的合成，但对原核生物没有作用。环己亚胺是 FKBP 及相关酶的竞争性抑制剂，同时也是细胞凋亡的诱导因素。

乙醇（75%和 95%）

据报道，有些批次的乙醇（100%或 200 proof）中含有会使微阵列试验中产生很深荧光背景的污染物质，95%的乙醇不存在这个问题。为避免这个问题，可以用 95%的乙醇或者已证明不会产生高荧光背景的无水乙醇。Warner-Graham 公司（Cockeysville，Maryland）出售的无水乙醇（USP 级乙醇，64-17-5）效果很好。

梯度缓冲液

150 mmol/L KCl

5 mmol/L $MgCl_2$

50 mmol/L Tris-Cl (pH 7.4)

高压灭菌后使用。

低渗裂解缓冲液

10 mmol/L KCl

1.5 mmol/L $MgCl_2$

10 mmol/L Tris-Cl (pH 7.4)

高压灭菌。使用前加入环己亚胺，使其终浓度达到 10 μg/ml。

异丙醇

磷酸盐缓冲液（PBS）

使用前加入环己亚胺，使其终浓度达到 10 μg/ml。

RNA 沉淀溶液

0.8 mol/L 柠檬酸钠

1.2 mol/L NaCl

醋酸钠（2 mol/L）

蔗糖缓冲液（1.95 mol/L）和蔗糖梯度缓冲液（1.3 mol/L）

用梯度缓冲液对 2.5 mol/L 蔗糖储存液做适当的稀释来制备这些溶液。每个梯度需要 13 ml 1.95 mol/L 和 6 ml 1.3 mol/L 的蔗糖缓冲液。在一种梯度缓冲液使用之前，加入环己亚胺使其终浓度为 10 μg/ml。

在制备含有真实裂解物的任何梯度缓冲液之前，最好制备一个模拟的梯度缓冲液并检查分层是否正确来测试这些溶液。

蔗糖梯度缓冲液（2.5 mol/L）

该溶液非常黏稠，配制时必须小心。

1. 以约 150 g 的增加量将 855.75 g 蔗糖加入到含有 250 ml、加热到 65℃的梯度缓冲液的 2 L 大烧杯中。
2. 在热台上用合适大小的搅拌子搅拌缓冲液，并将溶液温度保持在 65℃左右。如果溶液变得过度黏稠，可加入少量梯度缓冲液（65℃）。
3. 当所有的蔗糖加入后，用梯度缓冲液（65℃）将最终体积调节至 1 L。如果可能，可用折射计测量溶液中蔗糖的浓度，并对溶液浓度作相应的调整。
4. 用于梯度前，加入环己亚胺使其终浓度为 10 μg/ml。每个梯度需要 15 ml 2.5 mol/L 的蔗糖缓冲液。

Tris-Cl（10 mmol/L，pH 7.4）

TRIzol LS（GIBCO）

TRIzol LS 是 TRIzol 的浓缩物，在单相溶液中含有苯酚和异硫氰酸胍。TRIzol LS 是从液态样品中分离 RNA 非常好的试剂。

135

离心机和转子

Sorvall 离心机或相当的仪器

SS-34 离心机转子

SW28 超速离心转子

超速离心机（Backman 或 Sorvall）

专用设备

滚珠匀浆器或 5～10 ml 玻璃 Dounce 匀浆器

移液器用带滤芯吸头

电子移液泵

血球计

微量离心管（2 ml）
Oak Ridge 聚丙烯离心管（50 ml）（Nalge Nunc International）
塑料移液管（2 ml、10 ml 和 25 ml）
针（25.4 mm，18 号）
用于组织培养的圆培养瓶或细颈瓶
分光光度计
Ultra-Clear 25 mm×89 mm 离心管（344058，Backman 或相当的）。

方法

制备组织培养细胞

1. 在适当条件下生长组织培养细胞。

   因为每次蔗糖梯度离心都需要大量的细胞（5×10^8个），建议在圆培养瓶中生长细胞或在细颈瓶中悬浮培养；细胞也可以在皮氏培养皿中单层生长。悬浮培养物生长的浓度不能大于 5×10^5个/ml，因为密度过高将导致蛋白质合成的下调，从而减少了多核糖体mRNA的产量。

2. 向细胞培养物中加入环己亚胺，使其终浓度达到 10 μg/ml，然后将其放回培养箱或温室中。在继续下一步骤前，将环己亚胺处理过的细胞在 37℃ 孵育 5～10 min。

   重要事项：从这里开始，所有步骤应在 4℃或冰上进行。

3. 收集细胞。

136

滚动或悬浮培养物：

   a. 将细胞悬液转移到 250 ml 的离心瓶中。
   b. 用数百毫升含有 10 μg/ml 环己亚胺的冰冷 PBS 短时间漂洗空培养瓶，并用刮铲收集任何黏附在瓶壁上的残余细胞。
   c. 4℃下以 1000 *g* 离心 10 min，收集细胞。
   d. 弃去上清液，用 125 ml（每 250 ml 离心瓶）含有 10 μg/ml 环己亚胺的冰冷 PBS 清洗细胞两次。

生长在皮氏培养皿上的单层细胞：

   a. 吸去培养基。
   b. 加入少量冰冷的 PBS（含有 10 μg/ml 环己酰亚胺），并用橡胶刮勺从培养皿上刮下细胞。
   c. 在一个 250 ml 的离心瓶中合并细胞，并在 4℃下以 1000 *g* 离心 10 min，收集细胞。
   d. 弃去上清液，用 125 ml（每 250 ml 离心瓶）含有 10 μg/ml 环己亚胺的冰冷 PBS 清洗细胞两次。

4. 通过离心回收细胞，将细胞团块重悬于 10 ml 含有 10 μg/ml 环己亚胺的冰冷 PBS 中。
5. 将所有细胞都移入一支离心管中，并用血球计记数。将含有 5×10^5 个细胞的等

份转移到一系列的 15 ml 或 50 ml 圆锥形离心管中。4℃下以 1000 *g* 离心 10 min，收集细胞。

到这一步可以停止操作并将细胞团块放入液氮。将团块在−80℃下保存。冷冻后，细胞团块可使用至少 6 个月。

低渗裂解和梯度建立

6. 以 2.5×10^8个/ml 的浓度将细胞团块重悬于冰冷的低渗裂解缓冲液中。让细胞在冰上膨胀 5～10 min。

   可以向裂解缓冲液和接下来的梯度组分中加入 RNase 抑制剂。由于细胞系和组织中 RNase 的活性是变化的，因而必须根据每一个具体情况来决定。

7. 在一个配合紧密的 Dounce 匀浆器中打击 10 次来匀浆细胞，或在滚珠匀浆器中通过 10 次匀浆细胞。如果需要，可将一小部分裂解物放置于−80℃作为细胞总 RNA 的来源。
8. 4℃下将匀浆以 2000 *g* 离心 2 min 除去细胞核和未破碎的细胞。团块非常柔软且大小与细胞类型有关。移开上清液并用冰冷的低渗裂解缓冲液将其体积调节至 2 ml。

137

9. 将 2 ml 已调节体积的上清液（步骤 8）加到含有 11 ml 2.5 mol/L 蔗糖梯度缓冲液的 50 ml 锥形管中，涡旋以充分混合溶液。
10. 在 Ultra-Clear 25 mm×89 mm 离心管中建立蔗糖的不连续梯度。
    a. 在 Ultra-Clear 25 mm×89 mm 离心管底部加入 4 ml 2.5 mol/L 的蔗糖梯度缓冲液。
    b. 小心将步骤 9 中的 13 ml 裂解产物加到 2.5 mol/L 蔗糖层的上面。
    c. 将 13 ml 1.95 mol/L 蔗糖梯度缓冲液和 6 ml 1.3 mol/L 的蔗糖梯度缓冲液加入到梯度中去。

    使用电子移液泵和一个 10 ml 的塑料移液管能够很容易地建立各梯度层。将泵速调节到最小，使蔗糖溶液缓慢地流出移液管。保持移液管尖头始终处于上方，但是距离梯度的弯月面最顶端非常近。分层时，尽量不要碰到最顶端的部分。形成的梯度在层与层之间由于不同蔗糖溶液折射率的差异而存在清晰可见的平面。

11. 在 SW28 超速离心转子上，以 90 000 *g* 将梯度离心 5～6 h。

    精确平衡离心管非常关键。如果使用偶数个梯度，用 1.3 mol/L 的蔗糖梯度缓冲液来平衡。如果使用奇数个梯度，就要另外准备一个由 2 ml 水代替步骤 8 中的匀浆的模拟梯度，并且用 1.3 mol/L 的蔗糖梯度缓冲液来平衡所有的梯度。

    离心后，在 1.95 mol/L 和 1.3 mol/L 蔗糖溶液的交界处应该能看到含有膜物质的一层物质。

12. 用 18 号针刺破离心管底部收集梯度，并按照每 1.5 ml 分部分收集，放入 2 ml 离心管中。为了将游离 RNA 污染的风险减至最小，分离含有膜物质部分时首先用 P1000 Pipeman 从离心管顶部除去 10 ml 的溶液。剩余的梯度（含有细胞溶质的部分）通过穿刺的办法收集。

    在收集梯度时，为了将离心管固定，仔细地用锋利的刀或者剃须刀片将 50 ml 聚丙烯锥形管切成两半，并将上半部分丢弃。接下来，将仍然留在试管底部的锥形部分的最

尖端切除。这个装置可以用标准的夹具和具有合适的直径能够合身地放下超速离心管的圆环架悬在工作台的上方。为了收集各部分，手工将装有开口 2 ml 微量离心管的微量离心管架滑动到刺穿的梯度试管的下方。

13. 使用分光光度计测量每个部分在 260 nm 处的吸收以确定核酸的存在。

    游离的核糖体和 mRNA 一般存在梯度的加样区域，而膜结合的核糖体和 mRNA 存在于 1.95 mol/L 和 1.3 mol/L 蔗糖溶液的交界处。

RNA 分离

14. 以 OD_{260} 值作为指示，分别合并对应于加样区（游离 RNA）及位于 1.95 mol/L和 1.3 mol/L 蔗糖缓冲液交界处（膜结合的 RNA）的组分。用 TRIzol LS 试剂从每个合并物中分离总 RNA。

    a. 将合并的组分转移到 50 ml 的 Oak Ridge 聚丙烯离心管中。每份合并的组分中加入 3 倍体积的 TRIzol LS。

    b. 用移液管吸打进行混合，并剧烈涡旋。

    c. 将试管在室温下竖立静置 5 min，然后按每 0.75 ml 的 TRIzol LS 加入 0.2 倍体积的氯仿。使用带电动移液泵的玻璃移液管或者 1 ml 带滤芯的微量移液吸头以防氯仿“爬升”到较小吸头的滤芯中。倒转密闭的试管，剧烈振荡 15～20 s。

    d. 将试管在室温下竖立静置 5 min，4℃下以 12 000 *g* 离心 15 min。

    e. 用 10 ml 移液管将水相转移至一新的聚丙烯管中。水相的体积约为步骤 a 中使用的 TRIzol LS 体积的 70%。按每毫升 TRIzol LS 加入 0.25 ml 的异丙醇。然后，按每毫升 TRIzol LS 加入 0.25 ml 的 RNA 沉淀溶液，混匀溶液。

    f. 将样品室温下保存 10 min，然后在 4℃ 下以 12 000 *g* 离心 10 min 回收 RNA。

    g. 吸去上清液，加入 75%的乙醇轻轻倒转清洗 RNA 团块，4℃下 12 000 *g* 离心 5 min。

    h. 吸去上清液，让 RNA 团块在空气中干燥 10～20 min。将 RNA 团块重悬于 400 μl 10 mmol/L Tris-HCl (pH 7.4)，并转移至 1.5 ml 的微量离心管中。

    i. 向微量离心管中的 RNA 中加入 60 μl 2 mol/L 的醋酸钠及 1 ml 95%乙醇进行另一次乙醇沉淀。在干冰上或−80℃冷冻直到冻结。融化后在 4℃下以最大速度离心 30 min。

    j. 吸去上清液，用 500 μl 75%乙醇清洗 RNA 团块。4℃下以最大速度离心 15 min。

    k. 除去上清液，让 RNA 团块在空气中干燥 10～20 min。将 RNA 团块重悬于 40～200 μl DEPC 处理的水中。选用 **RNA 的保存和回收**信息栏中的一种方法储存 RNA。

        由于来源不同，两部分 RNA 的产量变化差异很大。膜结合的多核糖体 RNA 的产量在 10 μg 到数百微克之间，而“游离”RNA 的产量通常要高出 10～150 倍。

        通常，投入 30～50 μg 的总 RNA 可以与微阵列很好的杂交。如果梯度中膜部分回

138

收的 RNA 少于这个量，需要利用体外转录的策略来扩增两部分 RNA（Van Gelder et al. 1990；Eberwine et al. 1992；Philips and Eberwine 1996；Wang et al. 2000）（见方案 15 和方案 16）。

139

方案 7

用批量制备色谱法分离 poly（A）$^+$RNA

Damares Monte[*†]，**Shauna Somervile**[*]［[*]华盛顿卡内基研究院，斯坦福，加利福尼亚，94305；[†]E. M. B. R. A. P. A.（巴西农业研究公司）遗传资源与生物技术研究所，巴西利亚，巴西］

所有真核生物的 mRNA 在其 3′端都带有同聚腺嘌呤残基区域，因此可以通过 oligo（dT）-纤维素亲和层析从细胞的总 RNA 中纯化。这种方法由 Edmonds 等（1971）以及 Aviv 和 Leder（1972）首先提出，利用了 poly（A）尾巴可以和共价连接在纤维素支持介质上的 oligo（dT）（长度通常为 18～30 个核苷酸）形成稳定的杂合体的能力。因为只有一些 dT-A 碱基对形成，mRNA 与 oligo（dT）只有在高离子强度的缓冲液下才会发生反应。当不具有多聚腺嘌呤的 RNA 分子从介质上洗去之后，用低离子强度的缓冲液使双链结构的稳定性降低，并将 mRNA 从树脂上释放出来。

本方案描述了一种修改的 oligo（dT）-纤维素色谱法（Celano et al. 1993），推荐用于三种情况下：当制备的总 RNA 含有污染物（如蛋白质和多糖），可能会堵塞柱子时；操作少量的（<50 μg）总 RNA 时；同时处理多个样品时。也可以通过使用商品化试剂盒，如来自 QIAGEN 的 Oligotex mRNA 试剂盒批量制备色谱来分离 poly（A）$^+$ RNA。与 RNA 工作相关产品的商业参考列于表 3-5 中。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

重要事项：本方案中所有试剂使用经 DEPC 处理的 H_2O 配制（见 **RNA 操作的特别注意事项**信息栏）。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

乙酸铵（10 mol/L）〈!〉

DEPC 处理的 H_2O，冰冷

乙醇（95%和 75%）

据报道，有些批次的乙醇（100%或 200 proof）中含有会使微阵列试验中产生很深荧光背景的污染物质，95%的乙醇不存在这个问题。为避免这个问题，可以用 95%的乙醇或者已证明不会产生高荧光背景的无水乙醇。Warner-Graham 公司（Cockeysville，Maryland）出售的无水乙醇（USP 级乙醇，64-17-5）效果很好。

表 3-5 商品的网络资源

来 源	网站地址	试剂盒/产品[a]
QIAGEN	www.qiagen.com	RNA 纯化和清除试剂盒
Millipore	www.millipore.com	清除
Ambion	www.ambion.com	RNA 制备试剂盒
GIBCO	www.lifetech.com	酶、试剂
Amersham Biosciences	www.apbiotech.com	酶、试剂、荧光染料

a 使用这些商品时，总是要按照生产商的说明操作。

NaCl (5 mol/L)

Na-TES

这是含有 0.5 mol/L NaCl 的 TES

Oligo (dT)$_{18\text{-}30}$-纤维素（GIBCO 或相当的）

室温下用 Na-TES 缓冲液清洗 III 型纤维素（结合能力为每克 58～100 OD_{260} 单位）3 次进行水合，并按照 100 mg/ml 浓度重悬在 Na-TES 中。

TES

10 mmol/L Tris-HCl (pH 7.5)

1 mmol/L EDTA (pH 8.0)

SDS（0.1%，m/V）〈!〉

核酸和寡聚核苷酸

真核生物总 RNA

按照本章方案 2～方案 4 所述方法制备。

专用设备

带速度控制的微量离心机

事先调节到 37℃、55℃及 65℃的水浴

方法

1. 将约 1 mg 的总 RNA 溶解于 DEPC 处理过的水中，然后转移到一支灭菌的微量离心管中。用 TES 缓冲液将体积调节到 600 μl。
2. 将 RNA 在 65℃加热 5 min 以消除二级结构，然后将微量离心管从水浴中取出并立即插入冰水浴中。
3. 加入 60 μl 5.0 mol/L NaCl，将管内物质涡旋混合。
4. 加入 600 μl（约 60 mg）水合的 oligo（dT)-纤维素。用一次性移液器吸头将树脂和 RNA 溶液混合。
5. 将树脂浆在 37℃温浴 10 min，其间偶尔倒转轻轻混合。
6. 室温下以最大速度将离心管离心 5 min 使 oligo(dT)-纤维素成团块。弃去上清液。
7. 将团块用 1 ml Na-TES 重悬，并按照步骤 6 所述离心。弃去上清液。

8. 将团块用 1 ml 预冷的 DEPC 处理的水重悬，再离心一次。弃去上清液。
9. 用 400 μl DEPC 处理的水重悬 oligo（dT）-纤维素团块，加热 5 min 至 55℃，以洗脱多聚腺苷酸 RNA。离心步骤同前，但保留上清液。
10. 重复步骤 9 两次，每次的上清液分开保存。
11. 通过在 260 nm 和 280 nm 的吸收光谱或使用 RiboGreen 染色后的荧光对每份上清液中的 mRNA 进行定量（见 **RNA 的定量**信息栏）。合并具有最高产量的部分，并加入 0.2 倍体积的 10 mol/L 乙酸铵和 2 倍体积的无水乙醇沉淀 poly $(A)^+$ RNA。离心回收 RNA，并将 RNA 团块溶于含有 RNase 抑制剂如 RNasin 的 DEPC 处理的水中（可选的）（关于 RNase 抑制剂的选择，见 Sambrook and Russell 2001）。

    mRNA 的数量在不同细胞类型之间的差别很大。作为经验，利用 oligo（dT）—纤维素分离的 poly $(A)^+$ RNA 的量占总 RNA 的 1%～4%。

142

方案 8

用 CsCl 溶液超速离心分离 RNA

在本方法中（Chirgwin et al. 1979），培养的细胞或组织在 4 mol/L 的异硫氰酸胍中匀浆，裂解物置于高密度的氯化铯层上方。因为 RNA 在 CsCl 中的浮力密度（1.8 g/ml）远大于其他的细胞组分，rRNA 和 mRNA 在超速离心过程中迁移到管底（Glisin et al. 1974；Ullrich et al. 1977）。只要不连续梯度没有过载，蛋白质就会保留在胍盐裂解液中，而 DNA 浮在 CsCl 层中。由于 Chirgwin 方法得到的 RNA 具有很高的质量和纯度，而且操作不需大量人力，在 20 世纪 80 年代初期，该方法成为分离完整的高分子质量 RNA 的标准方法。尽管以后该方法被 Chomczynski 和 Sacchi（1987）提出的一步法所取代，但仍然有很多情况下优先采用 Chirgwin 的技术，例如：

- 从富含甘油三酯的脂肪组织中抽提 RNA 时（Tavangar et al. 1990）。
- 从富含多糖和（或）蛋白多糖的细胞或组织中抽提 RNA 时。据报道，这些污染物会阻碍乙醇沉淀后 RNA 的溶解，并抑制反转录反应（Groppe and Morse 1993；Schick and Eras 1995）。

下面介绍的方法是由 Nadia Traficante（Peter MacCallum Cancer Institute，Melbourne，Australia）提供的方案。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。
使用前将储存液稀释到合适的浓度。

CsCl (5.7 mol/L)/0.01 mol/L EDTA (pH 7.5)

CsCl/EDTA 溶液按照每批 100 ml 配制，将 96.0 g CsCl 溶解于 90 ml 0.01 mol/L EDTA (pH 7.5) 中，并加入 DEPC 使最终浓度为 0.1%。将溶液静置 30 min 后，在 15 psi (1.05 kg/cm^2) 下高压灭菌 20 min。溶液冷却后，用经 DEPC 处理的水调节体积至 100 ml。

EDTA (10 mmol/L，pH 8.0)

乙醇 (75%)

异硫氰酸胍匀浆缓冲液

4 mol/L 异硫酸氰胍 (M_r = 118.1)

0.1 mol/L Tris-Cl (pH 7.5)

1% β-巯基乙醇〈!〉

将 50 g 异硫氰酸胍溶解于 10 ml 1 mol/L Tris-Cl (pH 7.5)，加水至 100 ml。用 Whatman 1 号滤纸或与之相当的滤纸过滤溶液。该溶液性质稳定，可在室温下长期储存。使用前加入 β-巯基乙醇至最终浓度为 1% (0.14 mol/L)。

NaOH (0.1 mol/L)〈!〉

十二烷基肌氨酸钠 (20%，*w/V*)

细胞与组织

单层培养细胞

单层生长的细胞应该用 PBS 清洗一次，沥干；或者立即使用；或者在液氮中速冻并储存于 −80℃。可以按照如下方法操作：

- 用铝箔做一个浅的盘子，将组织培养板（盖好盖子的）排放在其中，不要将培养板堆叠放置。
- 缓缓向盘内注入液氮，使它在培养板的底部流动。当培养板内残存的液体冻结后，迅速用铝箔将培养板包好，储存于 −80℃。

储存在 −80℃ 冻结的单层细胞应于室温下直接在异硫氰酸胍匀浆缓冲液中融化。冻存的组织和细胞团块应该在液氮中用经过液氮预冷的研钵和研杵磨成粉末。使用一个较大的研杵以防止液氮加入研钵时粉末“爆炸飞散”。

悬浮培养细胞

悬浮培养的细胞要通过离心回收，用 PBS 清洗一次，再次离心，沥干，或者立即使用；或者在液氮中速冻并储存于 −80℃。

组织

组织应该是新鲜采集的，切成小块 (100 mg)，直接放入异硫氰酸胍匀浆缓冲液中，或者如不立即使用，在液氮中速冻并储存于 −80℃。

专用设备

Beckman SW41 转子（或相当的）和异质同晶聚合物 (polyallomer) 超速离心管

Polytron 匀浆器

配备 Probe 9100071 (Brinkmann 2713 700-8) 的 Polytron PT1200 匀浆器 (Kinematica AG, Switzerland) 或相当的仪器用来捣碎细胞和组织。使用前，清洁 Polytron 使其没有 RNase。拆开探针，先用水洗涤部件，然后用无水乙醇，最后用可降解 RNA 的 RNaseZap 处理。组

装探针并在水中匀浆。可以使用其他相当的仪器（FastPrep 仪器）。关于各种仪器的讨论，见**动物细胞和组织的匀浆**信息栏。

解剖刀片和止血钳

Sorvall 离心机和 SS-34 转子（或相当的设备）

方法

1. 组织粉末：室温下，将组织粉末加入含有异硫氰酸胍匀浆缓冲液的聚丙烯离心管中。1 g 组织粉末需要 10～20 ml 匀浆缓冲液。

   液体加得过多要比过少好，但是要考虑到转子的容量与样品的数量。一个 SW41 转子可容纳 6 支离心管、总体积约 42 ml 的匀浆。

   警告：操作装有缓冲液和组织粉末的聚丙烯试管时要十分小心。溶解氮的释放可能会冲开盖子，因而涉及匀浆缓冲液操作的所有阶段都要佩戴防护眼镜。

   培养细胞的裂解物：将裂解物转移到聚丙烯试管中。

2. 将裂解物在 Polytron 匀浆器中高速匀浆 1～2 min。加入十二烷基肌氨酸钠至最终浓度为 0.5%，并将悬液混合均匀。

   在 Polytron 匀浆器中匀浆剪切了核 DNA，并可以防止在 CsCl 层的顶部形成一个阻止 RNA 沉淀到离心管底部的不渗透层，匀浆能使 RNA 的产量提高约 2 倍。

3. 将悬液在室温下以 5000 *g* 离心 10 min。重复离心步骤直到所有漂浮的碎片都从悬液中除去。
4. 在 0.1 mol/L 的 NaOH 中清洗异质同晶聚合物超速离心管，并用水彻底漂洗，甩干离心管。向每支离心管中加入 5 ml 的 5.7 mol/L CsCl，10 mmol/L EDTA，然后小心将步骤 3 中的上清液加到 CsCl 层上。用防水记号笔在管外侧标记上顶层的位置。如有必要，用异硫氰酸胍匀浆缓冲液平衡离心管的重量。将不连续梯度在 20℃下以 150 000 *g* 离心 18 h。
5. 从离心机中取出试管时小心不要影响梯度。在每个试管距离底部约 0.5 cm 处划一条线。使用真空吸出器除去上清液直到 CsCl 层（试管外侧上方的线）。再用一个新的移液管取出直到试管外侧下方标记线的液体。
6. 用止血钳夹住一片烧红的解剖刀片，将离心管底部切开。不要掉下来！将试管底部倒置于一叠 Kimwipes 纸上。试管底部的 RNA 团块应该清晰可见并呈凝胶状。
7. 室温下向试管底部加入 70%乙醇。将试管残留部分倒置，沥干乙醇，确保团块的位置没有移动。让 RNA 团块在室温下晾干，然后用一支 P200 移液器上下吸打，重新溶解于约 200 μl 的 1 mmol/L EDTA（pH 8.0）中。
8. 将 10 μl 的 RNA 移到一个新的微量离心管中，加入 400 μl 1 mmol/L EDTA（pH 8.0）。用 1 mmol/L EDTA（pH 8.0）作参比，测量 RNA 溶液在 260 nm 和 280 nm 处的吸光值。计算 RNA 的产量。
9. 选用 **RNA 的保存和回收**信息栏中的一种方法储存 RNA。

方案 9

酵母总 RNA 的纯化

Charles B. Epstein（安万特制药公司，剑桥，马塞诸塞州）；
Ron Butow（得克萨斯大学，西南医学中心，达拉斯，得克萨斯）

用热的酸性苯酚（热酚法）处理酵母来实现 RNA 的分离。这种抽提方法能得到高分子质量的 RNA 而几乎没有 DNA 的污染。本方案是根据 Kohrer 和 Domdey（1991）的方法，用首先用酸性苯酚，然后用氯仿抽提从水相去除 DNA 及其他细胞组分。然而，如同**替代方案：用锁相胶纯化 RNA** 的信息栏中描述的那样，水相的分离与回收以及残留污染物的去除也可以通过使用 Phase Lock Gels 很容易地实现。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

重要事项：本方案中所有试剂使用经 DEPC 处理的 H_2O 配制（见 **RNA 操作的特别注意事项**信息栏）。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

酸性苯酚（pH 5.0～5.1）〈!〉

为了得到酸性苯酚，将一瓶（100 g）新鲜的结晶苯酚加热到 65℃。液化的苯酚呈透明无色状，呈黄色或粉红色的苯酚批次应弃去不用。加入 20 ml 的 10 mmol/L 醋酸钠（pH 5.1）、50 mmol/L NaCl 和 1.0 mmol/L EDTA（pH 8.0）混合液，搅拌混合瓶内组分。将液化的苯酚冷却到室温，再加入 10 ml 缓冲液。用铝箔包裹瓶体并在 4℃保存。要测试苯酚的 pH，参考用苯酚抽提蛋白和 DNA 的信息栏中的用法说明。

AE 缓冲液

50 mmol/L 醋酸钠（pH 6.0）

10 mmol/L EDTA

醋酸铵（7.5 mol/L）

氯仿：异戊醇（24：1）〈!〉

DEPC 处理的水

乙醇（无水）

据报道，有些批次的乙醇（100%或 200 标准强度）中含有会使微阵列试验中产生很深荧光背景的污染物质，95%的乙醇不存在这个问题。为避免这个问题，可以用 95%的乙醇或者已证明不会产生高荧光背景的无水乙醇。Warner-Graham 公司（Cockeysville，Maryland）出售的无水乙醇（USP 级乙醇，64-17-5）效果很好。

146 异丙醇

苯酚：氯仿〈!〉

50%液化苯酚

50%氯仿，用 AE 缓冲液平衡后于 4℃保存在避光瓶中。

SDS（10%，*m/V*）

醋酸钠（3 mol/L，pH 5.2）

培养基

YPD 或其他适合的酵母生长培养基

细胞和组织

酵母细胞，作为过夜培养物新鲜生长

离心机和转子

能容纳 250 ml 瓶子的离心机

Oak Ridge 试管（耐苯酚的）

Sorvall 离心机，SS-34 和 GSA 转子（或与之相当的）

专用设备

65℃水浴

方法

酵母收集与 RNA 抽提

1. 接种并将酵母培养物（250 ml）在合适的条件下生长至 A_{600} 为 0.8。

   标准的酵母生长条件是 30℃下在振动培养箱中的 YPD 培养基上生长。将大约 25 μl 野生型酵母菌株接种到 250 ml 的 YPD 培养基上过夜培养，接种 16 h 后，培养物 A_{600} 约为 0.72。为了从指数生长的细胞中分离 RNA，培养物的 A_{600} 应为 0.5～1.0。

2. 室温下以 2500 *g* 离心 5 min，收集细胞。弃去上清。

   在这个阶段，细胞可作为冰冻的团块保存于液氮中或者立即用于 RNA 的抽提和纯化。在室温下离心是为了防止冷休克，这将会改变细胞的转录模式。

3. 将细胞重悬于 10 ml AE 缓冲液中，再将细胞悬液转移到 50 ml 的聚丙烯离心管中。

4. 加入 1 ml 10% SDS 和 11 ml 酸性酚保持在 65℃。涡旋 15～20 s 混合管内组分。将试管转移到 65℃水浴中，放置 1 h，每隔 10 min 用最大速度涡旋 10 s。

147

5. 最后一次涡旋后，将试管在冰水浴中搅动，冷却到室温。

   不要使乳液冷却到 15℃以下，否则 SDS 会沉淀出。

6. 室温下以 2500 *g* 离心 5 min 进行相分离。

7. 用移液管移走并丢弃 10 ml 的下层（有机）相，将界面和细胞碎片保留在管中。

8. 向试管中加入 10 ml 热的酸性酚，并转移到 65℃水浴中 15 min，每隔 5 min 用最大速度涡旋 15 s。
9. 将乳液冷却到室温，并在室温下以 2500 *g* 离心 5 min 进行相分离。
10. 将上层水相（约 11 ml）转移到含有 11 ml 氯仿：异戊醇（24：1）的新的聚丙烯管中。盖上管盖剧烈涡旋 30 s，混合管内组分，然后以 2500 *g* 离心 5 min 进行相分离。

RNA 纯化

11. 将上层水相（约 11 ml）转移到新的聚丙烯管中。加入 2.6 倍体积的含有 0.3 mol/L 醋酸钠的无水乙醇。充分混匀，在－20℃保存至少 6 h。
12. 以 3500 *g* 离心至少 20 min 来回收 RNA。立即小心地将含有乙醇的上清液转移到新的离心管中，在－20℃保存直到方案完成，并且得到满意的 RNA 产量。
13. 加入 15 ml 冰冷的 75%乙醇，短暂涡旋混合然后离心。立即小心地将含有乙醇的上清液转移到新的离心管中，在－20℃保存直到方案完成，并且得到满意的 RNA 产量。
14. 小心地从管壁吸走最后一滴乙醇，并让 RNA 团块在室温下干燥。
15. 将几乎不湿润的 RNA 团块用 0.5～1.0 ml 10 mmol/L Tris（pH 8.0）溶解。将溶液转移到微量离心管中，65℃加热 5 min 以促进 RNA 的溶解。
16. 将试管在室温下以 12 000 *g* 离心 2 min。将上清液按照 250 μl 的等份转移到新的微量离心管中。将含有团块的微量离心管弃去，沉淀中主要包含大量的糖类及细胞壁组分。
17. 加入 2.5 倍体积的含有 0.3 mol/L 醋酸钠的无水乙醇。试管在－20℃保存至少 1 h，然后在微量离心管中以 12 000 *g* 离心 15 min，回收 RNA。
18. 吸出上清液，让 RNA 团块在室温下干燥。将每个团块溶解于大约 200 μl 10 mmol/L Tris（pH 8.0）。如果团块不能立即溶解，将溶液在 45℃温浴 10 min，然后在 4℃保存过夜。
19. 在微量离心机中以 12 000 *g* 离心除去任何颗粒状物质。合并上清液，通过在 260 nm 和 280 nm 的吸收光谱或使用 RiboGreen 染色后的荧光对每份上清液中的 mRNA 进行定量（见 **RNA 的定量**信息栏）。
20. 选用 **RNA 的保存和回收**信息栏中的一种方法储存 RNA。

148

将 RNA 用于微阵列实验前，通过甲醛-1%琼脂糖凝胶电泳检查 RNA 的完整性（见**哺乳动物、植物和细菌 RNA** 的信息栏）。凝胶中应该显示明显的 rRNA 和低分子质量的 RNA 条带，叠加了非常浅的扩散的 mRNA。

250 ml 健康的野生型酵母菌株生长到 OD_{600} 为 0.5～1.0 的培养物可得到 5～10 mg 的总 RNA。

替代方案：用锁相胶纯化 RNA

生产商提供的装在试管中的锁相胶在（phase lock gels）离心力作用下发生迁移，在有机相和水相之间形成一层屏障。界面处和下层的有机物质与上方的水层被有效地分离。这种胶是防漏的，能使上清液水相倒入新的试管时不会明显地带入来自界面中的物质污染。锁相胶有多种配方，分装在不同大小的试管中。关于详细信息，请参考生产商的资料或者 www. brinkman. com。

附加材料

锁相胶试管（50 ml）（955 15 409-6，Brinkmann Eppendorf）
台式离心机

方法

1. 进行主要方案的步骤 1～6。
2. 以 1500 *g* 离心 1 min 使 50 ml 试管中的锁相胶呈团块状。将步骤 6 中的上清液加入到含有锁相胶的试管中。加入 15 ml 氯仿并振荡试管使不同的相完全混合。不能涡旋振荡试管!
3. 室温下在台式离心机中以 1500 *g* 离心 5 min 使之分相。
4. 仔细将上层水相轻轻倒入一支新的 Oak Ridge 管中。
5. 测量溶液体积，加入 0. 1 倍体积的 3 mol/L 醋酸钠（pH 5. 2）和 1. 1 倍体积的异丙醇，室温放置 10 min。
6. 于 15℃下，在 Sorvall SS-34 转子中以 16 000 *g* 离心 5 min，回收 RNA 沉淀。
7. 轻轻吸去上清液，并用 70％乙醇清洗 RNA 团块和试管侧壁。再次于 15℃下，在 Sorvall SS-34 转子中以 16 000 *g* 离心 20 min。
8. 进行主要方案的步骤 15～20。

149

方案 10

E. coli RNA 的纯化

Virgil Rhodius（加州大学，旧金山，加利福尼亚，94143）

用热的酸性苯酚（热酚法）处理酵母来实现 RNA 的分离。首先用酸性苯酚，然后用氯仿抽提从水相去除 DNA 及其他细胞组分。要在 RNA 反转录之前用 DNase I 进行后续处理。这种抽提方法得到高分子质量的 RNA，且几乎没有 DNA 的污染。由于生长和收集的条件不同，从 10 ml 培养物得到的产量为 70～300 μg。该方案根据 www. microarray. org/protocols. html 和冷泉港实验室微阵列课程（cold spring harbor mi-

croarray courses）使用的技术改编。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录3。

重要事项：本方案中所有试剂使用经DEPC处理的H_2O配制（见**RNA操作的特别注意事项**信息栏）。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录2。

使用前将储存液稀释到合适的浓度。

氯仿〈!〉

乙醇（无水和80%）

据报道，有些批次的乙醇（100%或200标准强度）中含有会使微阵列试验中产生很深荧光背景的污染物质，95%的乙醇不存在这个问题。为避免这个问题，可以用95%的乙醇或者已证明不会产生高荧光背景的无水乙醇。Warner-Graham公司（Cockeysville，Maryland）出售的无水乙醇（USP级乙醇，64-17-5）效果很好。

乙醇/苯酚停止溶液〈!〉

水饱和的苯酚（pH<7.0），在乙醇中（5%，*V/V*）。

该溶液的配方由Jon Berstein（University of California，San Francisco）提供。

水饱和的苯酚（pH<7.0）

用Ultra-Pure苯酚（重蒸的，晶体；15509，Invitrogen/GIBCO）制备苯酚溶液。注意水饱和的酸性酚用于RNA抽提。

酚：氯仿（1∶1，*V/V*）〈!〉

SDS（10%和5%，*w/V*）〈!〉

醋酸钠（1 mol/L和3 mol/L，pH 5.2）

Tris（10 mmol/L，pH 8.0）

酶和缓冲液

150

溶菌酶缓冲液

500 μg/ml，在10 mmol/L Tris（pH 8.0）中

1 mmol/L EDTA（pH 8.0）

溶菌酶缓冲液应在使用之前新鲜配制。

5×DNase I缓冲液

50 mmol/L $MgCl_2$

50 mmol/L Tris-Cl（pH 7.5）

5 mmol/L EDTA（pH 8.0）

5 mmol/L 二硫苏糖醇（DTT）

不含RNase的DNase（10单位/μl）（776785，Boehringer Mannheim）

RNase抑制剂（799017，Boehringer Mannheim）

培养基

细菌培养基，适合 *E. coli* 菌株生长

细菌菌株

选择用于 RNA 制备的 *E. coli* 菌株

专用设备

Sorvall RC5C，SS-34 转子（或与之相当的）
64℃水浴

方法

总 RNA 的收集

1. 向培养物接种合适的 *E. coli* 菌株，在 37℃过夜生长达到饱和。
2. 在 100 ml 合适的培养基中稀释 1 ml 过夜生长物，使培养物生长到对数中期。
3. 将 10 ml 培养物转移到含有 1.25 ml 冰冷的乙醇/氯仿终止溶液的 15 ml 锥形试管中。

   本步骤终止 RNA 的降解。

4. 在 4℃下以 6000 *g* 离心 2 min 回收细胞。吸去培养基。

   在这个阶段，细胞可以按要求在液氮中速冻后保存于−80℃或立即用于 RNA 的抽提。

5. 用 500 μl 的溶菌酶缓冲液重悬细胞团块，加入更多的溶菌酶缓冲液将体积调节到 800 μl。将裂解产物转移到含有 80 μl 10% SDS 的 2 ml 微量离心管中。用移液器吸头上下吹打混匀组分。将细胞裂解物在 64℃下孵育 2 min。

   在孵育过程中，裂解物应该是澄清的。

6. 将 88 μl 1 mol/L 醋酸钠溶液(pH 5.2)加入到澄清的裂解物中，并倒转试管混合。

151

热苯酚抽提 RNA

7. 向裂解物中加入等体积（约 1 ml）水饱和的酚。盖上管盖，倒转试管 10～12 次混合组分。将试管在 64℃水浴孵育 6 min。在孵育过程中每隔 40～60 s 混合管内组分一次。
8. 将试管转移到冰浴中放置 2 min。4℃下在微量离心机中以最大速度（16 000 *g*）将乳液离心 10 min。
9. 将上（水）层转移到新的 2 ml 微量离心管中，然后加入等体积的氯仿。盖上管盖，倒转试管 8～10 次混合组分。
10. 在 4℃下以最大速度（16 000 *g*）在微量离心机中离心 5 min 使其分相。
11. 回收上层（水）相。将溶液等分到两个 1.5 ml 的微量离心管中，加入 0.1 倍体积的 3 mol/L 醋酸钠（pH 5.2）和 2.5 倍体积的冷乙醇（无水）。将溶液在

−20℃放置至少 1 h。

12. 4℃下，在微量离心机中以 12 000 *g* 离心 30 min，回收 RNA。
13. 用 80%冰冷的乙醇清洗 RNA 团块。4℃下，在微量离心机中以最大速度(16 000 *g*)离心 5 min。小心吸去乙醇溶液，并将开口的试管竖放在化学通风橱中直至最后一丝乙醇蒸发掉。
14. 将每个 RNA 团块溶解于大约 100 μl 水中。合并一对试管中的 RNA 溶液。

DNase 酶处理

15. 向每个 RNA 制备物（200 μl）中加入：

RNase 抑制剂（20 单位）	0.5 μl
DNase I 缓冲液	50 μl
不含 RNase 的 DNase（10 单位/μl）	1.0 μl

将反应在 37℃孵育 30 min。

16. 用等体积水饱和的酚（pH<7.0）抽提溶液。将乳液在微量离心机中以最大速度（16 000 *g*）离心 2 min。将上层（水）相转移到新的微量离心管中。
17. 用等体积的酚：氯仿抽提溶液。倒转试管数次进行混合，将乳液在微量离心机中以最大速度（16 000 *g*）离心 2 min。将上层（水）相转移到新的微量离心管中。

    一些 *E. coli* 菌株可能需要额外的苯酚和（或）氯仿抽提来去除污染的 RNase。在继续操作之前，建议进行两次额外的氯仿抽提（通过重复步骤 9 和步骤 10）。抽提后，将水相转移到新的微量离心管中，并继续步骤 18。

18. 加入 0.1 倍体积的 3 mol/L 乙酸钠（pH 5.2）和 2.5 倍体积的乙醇。将溶液在−20℃放置至少 1h 或者过夜。
19. 4℃下，在微量离心机中以最大速度（16 000 *g*）离心 30 min，回收 RNA。

152

20. 用 1 ml 冰冷的 80%乙醇清洗 RNA 团块。4℃下在微量离心机中以最大速度（16 000 *g*）离心 5 min。小心吸去乙醇溶液，并将开口的试管竖放在化学通风橱中直至最后一丝的乙醇蒸发掉（约 15 min）。

鉴定 RNA 的产量和质量

21. 将每个 RNA 团块溶解于大约 50 μl 的 10 mmol/L Tris (pH 8.0)。
22. 通过在 260 nm 和 280 nm 的吸收光谱或使用 RiboGreen 染色后的荧光对 RNA 进行定量（见 **RNA 的定量**信息栏）。

    根据 *E. coli* 菌株、生长条件和收集时培养物的密度不同，RNA 的产量应为 70～300 μg。

23. 选用 **RNA 的保存和回收**信息栏中的一种方法储存 RNA。

    E. coli RNA 在−20℃可以保存数月。对于更长时间的保存，可将溶液用乙醇沉淀，并在−80℃存放。

    将 RNA 用于微阵列实验前，通过甲醛-1%琼脂糖凝胶电泳检查 RNA 的完整性（见**哺乳动物、植物和细菌 RNA** 的信息栏）。凝胶中应该显示明显的 23S 和 16S rRNA 和低分子质量的 RNA 条带。23S 条带的染色要比 16S 条带高出两倍以上。严重的扩散或者低分子质量 RNA 染色很强表明了 RNA 可能降解或有 DNA 污染。

153

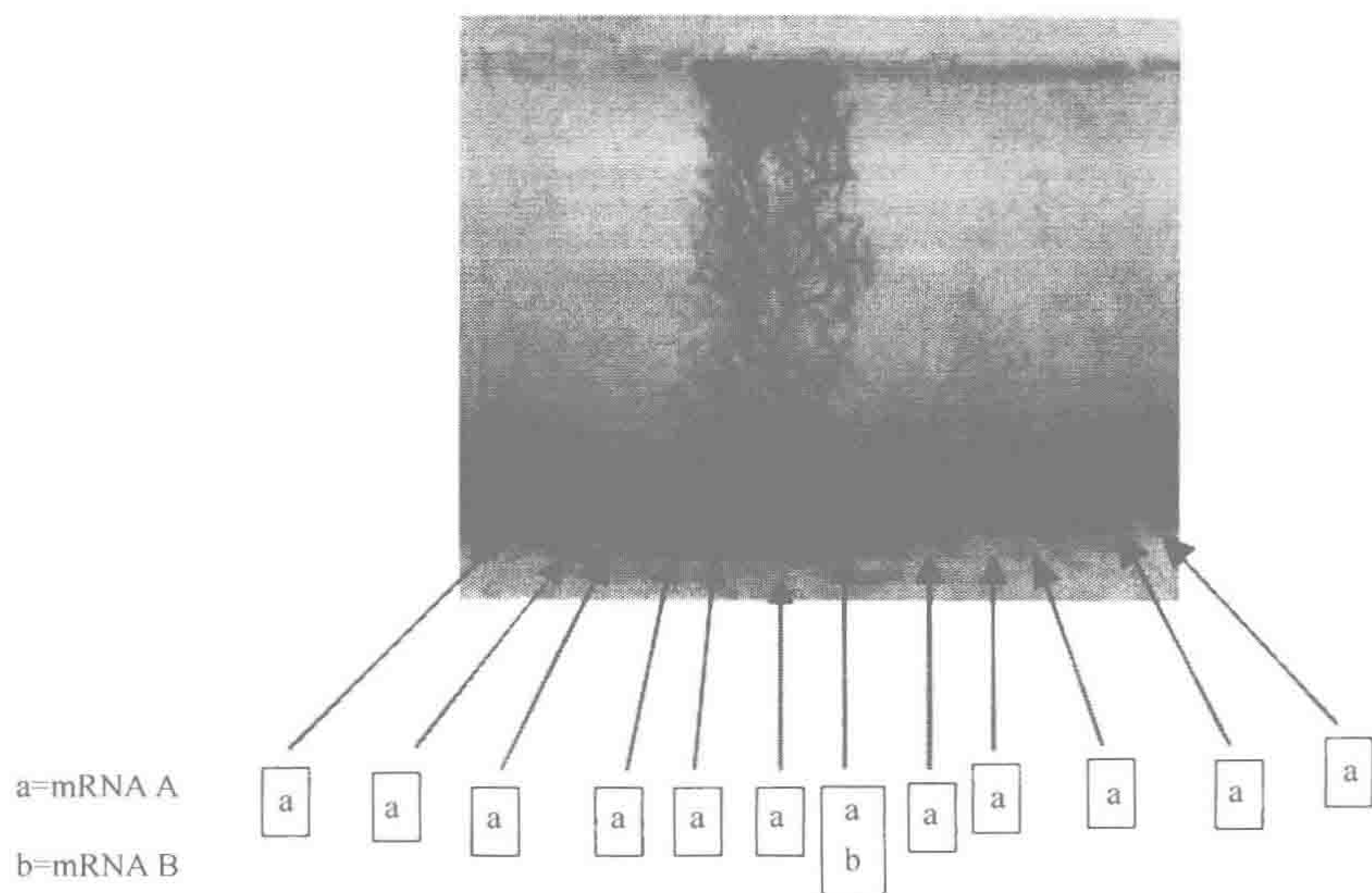
方案 11

从单细胞和少量组织中分离和扩增 RNA

James Eberwine,[*†] Paolo Marciano,[‡] Tracy McIntosh,[‡] Janet Estee Kacharmina[*†]（[*]药理学，[†]精神病学，[‡]神经外科学系，宾州大学医学中心，费城，宾夕法尼亚，19104）

使用 DNA 微阵列进行基因表达分析的标准方案是为从数百万个新鲜收获的细胞或者数毫克新鲜切割的组织中分离的大量 RNA 而研究和优化的。由于几乎没有组织仅含有一种细胞类型，所以表达谱分析需要的大规模 mRNA 制备物在大多数情况下是从复杂的和不同种类的混合类型细胞中分离得到的。这导致对一个特定基因表达的时间和空间所作的估计从最佳角度来看是不严密的。

神经组织为这个问题提供了一个图示说明。在中枢神经系统中，大脑的海马区是与学习和记忆有关的。海马由多种类型的细胞组成，包括各种神经元和神经胶质细胞。如果想要研究与学习挑战有关的基因表达谱变化，对从整个海马得到的 mRNA 分析必定是一无所获。在三维空间上的复杂性和特定的细胞反应都会丢失。在理想情况下，最好用所选择的感兴趣的细胞，甚至是单个细胞的 mRNA 进行分析（图 3-5）。正如下面的方案中所述，RNA 的扩增为实现这些目标和从单个细胞（这里是固定的神经元）中产



154

图 3-5　细胞稀释效应

图中显示了大鼠海马的 CA1 区域。单个锥体细胞用生物胞素标记来显示细胞的各个方面。CA1 区域中排成行的细胞用箭头突出表示。如果整个 CA1 区域被收集，来自生物胞素标记的细胞的 mRNA 被来自其他细胞的 RNA 所稀释。如果一种特定的 mRNA A 存在于所有的细胞中，那么应用微阵列分析它将很容易检测到；然而，如果 mRNA B 仅存在于生物胞素标记的细胞中，那么它的浓度至少会被稀释 10 倍，使得对这个单个细胞的全体 mRNA 进行定量会更加困难

生标记的靶分子提供了一个途径。

从固定的组织样品中分离的 mRNA 分析起来很困难，因为固定剂如戊二醛、福尔马林和多聚甲醛在伯胺之间形成交联。用这些试剂固定的组织中，核酸和蛋白质之间互相形成共价耦联。结果，从固定组织分离的 RNA 通常是片段化的，因为多聚核苷酸链与它们的束缚物分离开会导致链的断裂。然而并不需要将 RNA 从固定的组织中抽提出来，以生成用于微阵列分析的靶分子。相反，组织切片中的 mRNA 能够直接在组织切片上原位拷贝为 cDNA。图 3-6 显示了从组织切片进行 RNA 的反转录和扩增所包含的步骤（见本方案最后**替代方案：组织切片中 RNA 的原位转录**信息栏）。

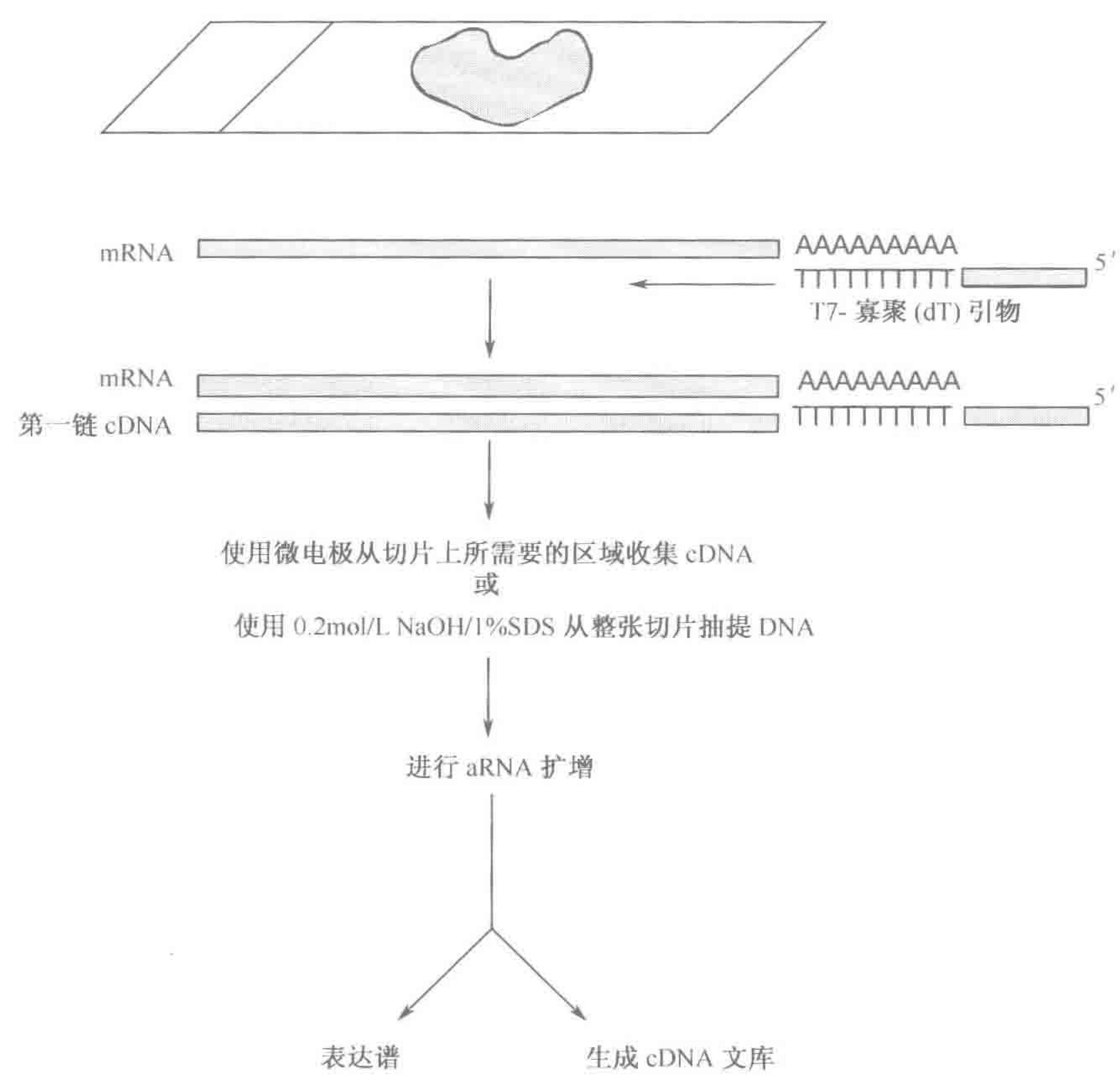


图 3-6　固定组织切片中的 cDNA 合成

在细胞培养物的固定组织切片上进行原位反转录反应。用带有噬菌体 T7 启动子的 oligo（dT）引物启动 cDNA 第一链的合成。可以使用微电极从单个的细胞来收集 cDNA，或者用碱/SDS 从整个切片中抽提

- 通过显微切割分离单个细胞，随后进行第一链的合成，以 oligo（dT)$_{24}$-T7 引物启动 cDNA 合成。
- 用一种 mRNA—特异性引物与组织切片原位杂交，通常是在常规的反义 RNA（aRNA）技术中使用的 oligo（dT)$_{24}$-T7 引物（主要方案）。引物的 oligo（dT）部分复性到组织切片中 mRNA 的 poly（A）尾部并提供了 3′-羟基末端，在常规的反转录反应中进行原位延伸。反转录酶拷贝 mRNA 模板，跳过交联的核苷，

产生相对较长的固定的 mRNA 的 cDNA 的拷贝。通过显微操作从组织切片中回收单个细胞，并抽提单链 cDNA。

- cDNA 第二链的合成。根据回收的第一链产物（来自单个细胞或者来自组织切片）在分子一个末端单链发夹结构处延伸的自我引发反应（self-priming reaction）。
- 用一种单链特异性核酸酶切开发夹环后，得到的双链 DNA 作为转录 aRNA 的模板。
- aRNA 再次转变为双链 DNA，在突出端补齐后，作为下一轮体外转录为 aRNA 的模板。到此阶段，RNA 的量已经扩增到可用作模板以产生用于 DNA 微阵列的荧光标记靶分子。

主要方案中描述了单个细胞的分离和随后的 RNA 扩增，而替代方案中提供了一种将 RNA 原位拷贝成 cDNA 的方法。一个附加方案（**附加方案：切片的吖啶橙染色**）描述了使用吖啶橙的细胞免疫染色来检测 RNA 的存在，并且据此来决定继续进行单细胞分离或者原位转录的可行性。

156

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

重要事项：本方案中所有试剂使用经 DEPC 处理的 H_2O 配制（见 **RNA 操作的特别注意事项**信息栏）。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度并过滤灭菌。

二硫苏糖醇（DTT）（100 mmol/L）〈!〉

乙醇（无水或者 95%）

据报道，有些批次的乙醇（100%或 200 标准强度）中含有会使微阵列试验中产生很深荧光背景的污染物质，95%的乙醇不存在这个问题。为避免这个问题，可以用 95%的乙醇或者已证明不会产生高荧光背景的无水乙醇。Warner-Graham 公司（Cockeysville，Maryland）出售的无水乙醇（USP 级乙醇，64-17-5）效果很好。

甲醛（37%）〈!〉

甲酰胺（去离子的）〈!〉

10×MOPS

酚：氯仿（1∶1，V/V）〈!〉

10×RNA 扩增缓冲液

400 mmol/L Tris-碱（pH 7.5）

70 mmol/L $MgCl_2$

100 mmol/L NaCl

20 mmol/L 亚精胺

加入亚精胺前先过滤

10×第二链缓冲液

1 mol/L Tris-碱（pH 7.4）

200 mmol/L KCl

100 mmol/L $MgCl_2$

400 mmol/L $(NH_4)_2SO_4$

分装成小份在－20℃储存。

醋酸钠（3 mol/L）或氯化钠（5 mol/L）

三氯乙酸（TCA）（10%，*m/V*）〈!〉

酶和缓冲液

AMV-RT（25 000 单位/ml，2000 单位）（禽成髓细胞瘤病毒反转录酶）（800-237-4512，Cape Cod Inc.，先前的 Seikagaku America）

AMV 是本方案中优先选择的酶，因为这个酶中含有微量的 RNase H 有助于引发第二链的合成。根据作者本人的经验，发现 AMV 比 SUPERSCRIPT II 能够产生更长的 cDNA 产物和更高的产量。

10× AMV-RT 缓冲液

500 mmol/L Tris-碱（pH 8.3）

1.2 mol/L KCl

100 mmol/L $MgCl_2$

E. coli DNA 聚合酶 I 的 Klenow 片段（2 单位/μl；标记级）（Roche，800-262-1640）

157

10× KFI（Klenow 补平反应）缓冲液

200 mmol/L Tris-碱（pH 7.5）

100 mmol/L KCl

50 mmol/L NaCl

50 mmol/L DTT

DTT 可以在缓冲液中略去而直接加入到反应中（因为时间过长会从缓冲液中沉淀出来）。

Nuclease S1（100 单位/μl）

10×Nuclease S1 缓冲液

2 mol/L NaCl

500 mmol/L 醋酸钠（pH 4.5）

10 mmol/L $ZnSO_4$

RNasin（40 单位/μl）

T4 DNA 聚合酶（1 单位/μl）（Roche，800-262-1640）

T7 RNA 聚合酶（1000 单位/μl）（Epicentre Technologies，800-284-8474）

核酸和寡核苷酸

载体 RNA（*E. coli* tRNA）

CTP（1.0 mmol/L）

[α-^{32}P] CTP（300 Ci/mmol；1 mCi/100 μl）（注：1Ci=3.7×10^{10} Bq）〈!〉
含四种 dNTP 的 dNTP 溶液，每种浓度为 2.5 mmol/L
含四种 NTP 的 NTP 溶液，每种浓度为 2.5 mmol/L
含有 ATP、GTP 和 UTP 的 3 NTP 溶液，每种浓度为 2.5 mmol/L
随机六聚体（Roche，800-262-1640）

1 A_{260}单位=40 μg

50 A_{260}=50×40 μg=2 mg

T7-Oligo $(dT)_{24}$寡聚核苷酸（100 ng/μl）

AAA CGA CGG CCA AAT TGT AAT ACG ACT CAC TAT AGG CGC $(T)_{24}$
关于该引物的描述见方案的介绍。

凝胶

琼脂糖凝胶（1%），含有 2.2 mol/L 甲醛〈!〉

详细情况见 Sambrook 和 Russell（2001）。

10×凝胶上样染料

细胞和组织

组织切片，固定到载玻片上

专用设备

电极固定器（3 个端口；Warner Instruments，Connecticut）
显微切割用微电极

微电极经过高压灭菌，并且在 Sutter P-97Puller（Sutter Instruments Co，California）上拉制为约 3 μm 开口。

Millipore 过滤器（孔径 0.025 μm）
Narishige 操作器，粗略的（Narishige internationa，New York）
Narishige 操作器，精细的（Narishige internationa，New York）
PE 管（直径 2 mm；Tygon）

方法

收获单个细胞

1. 向组织切片上加入约 200 μl DEPC 处理过的 ddH_2O 使之水合。
2. 向用于显微切割的微电极的尾部注入 20 μl DEPC 处理过的 ddH_2O，然后将它固定在 3 个端口的电极固定器上。
3. 将 PE 管的一端连接到 Warner 电极固定器的第三个端口上，并将固定器和显微镜固定形成一个牢固的吸引装置。将管子的另一端连接到 1 ml 的注射器上。
4. 用一个粗略的 Narishige 操作器将电极下降到 DEPC 处理的 ddH_2O 中，然后用精细的 Narishige 操作器将电极下降到视野内。

5. 当电极和组织都在视野内时，将电极从一侧移动到另一侧，然后降低电极直到它触及感兴趣的细胞。一旦电极处于合适的位置，通过将细胞膜从周围组织分离来缓慢地显微切割细胞。

    这个过程序必须慢慢地仔细操作，以避免破坏相邻的组织或者其他来自细胞的 mRNA 污染电极。

6. 一旦细胞与细胞外基质分离，将抽气泵连到 1 ml 的注射器上把细胞吸入电极。用粗略的操作器从组织上提升含有细胞的电极（关于单个细胞选择之前、期间和以后的示例，见图 3-7）。
7. 将电极内的物质排出到一个微量离心管中。

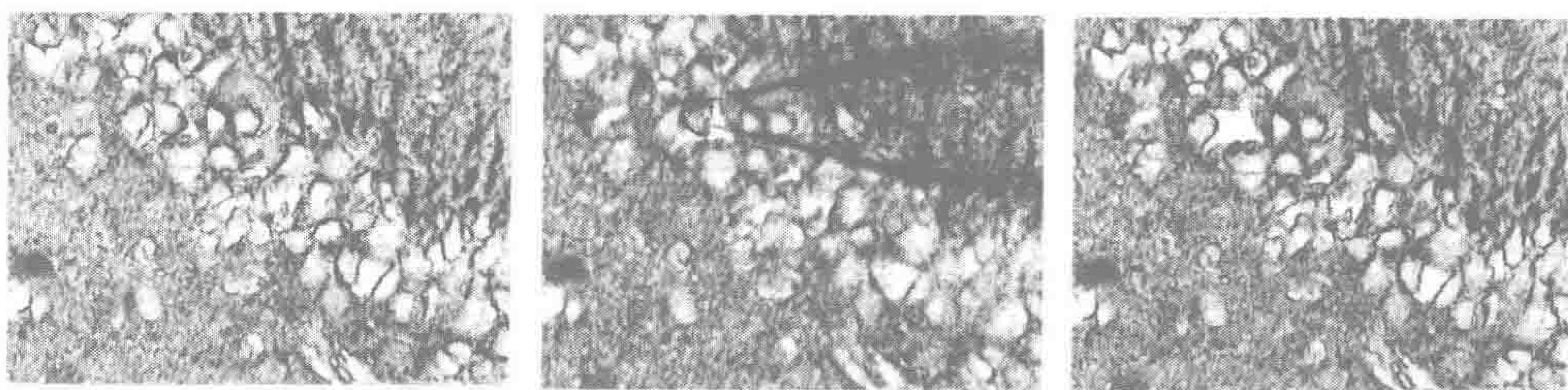


图 3-7　机械方式收集单个细胞

（左）原位转录后小鼠大脑的切片，其中可见各种细胞，但是在收集之前；（中）将膜片吸管叠放在要收集的细胞上；（右）细胞收集后的切片。细胞移去后留下的空洞很容易看出来

cDNA 的合成

8. 准备含有下列试剂的反应混合物：

DEPC 处理的 ddH_2O	4.5 μl
dNTP 溶液（每种的浓度为 2.5 mmol/L）	4.0 μl
10×RT 缓冲液	4.0 μl
T7-oligo $(dT)_{24}$（100 ng/μl）	2.5 μl
DTT（100 mmol/L）	1.0 μl
RNasin（40 单位/μl）	1.0 μl
AMV-RT	3.0 μl

将步骤 7 中试管内的物质转移到反应混合物中并在 37℃下将反应孵育 60 min。

9. 用等体积的酚：氯仿抽提混合物然后向水相中加入 1/10 体积的盐（3 mol/L NaAcO 或 5 mol/L NaCl）。
10. 向水相中加入 2.5 倍体积的 95%乙醇和 0.5 μg *E. coli* tRNA。在 4℃静置 30 min 来沉淀样品；4℃下在微量离心机中以最大离心速度（16 000 *g*）离心 20 min 回收核酸。
11. 除去乙醇，立即倒转试管并将团块在空气中干燥 10 min。将团块重悬于 20 μl DEPC 处理的 ddH_2O 中。

第二链 cDNA 合成

12. 为了使 RNA-DNA 杂交体变性，将样品在 85℃加热 5 min，然后立即放在冰上。
13. 向变性的样品中加入下列试剂：

DEPC 处理的 ddH_2O	10 μl
10×第二链缓冲液	4 μl
dNTP 溶液（每种的浓度为 2.5 mmol/L）	4 μl
T4 DNA 聚合酶（1 单位/μl）	1 μl
Klenow 片段（2 单位/μl）	1 μl

在 14℃下将反应孵育 4 h 至过夜。

这是一个自我引发的反应，cDNA 第一链的 3′端用作引发第二链的合成。

14. 为使环的切开容易进行，加入下列试剂：

DEPC 处理的 ddH_2O	319 μl
10×核酸酶 S1 缓冲液	40 μl
核酸酶 S1（在 S1 缓冲液中稀释为 1 单位/μl）	1 μl

在 37℃下将反应孵育 5 min。

15. 用等体积的（400 μl）酚：氯仿抽提反应。
16. 加入 1 ml 冰冷的乙醇并将溶液在－80℃孵育 30 min 来沉淀样品。4℃下在微量离心机中以最大离心速度（16 000 *g*）离心 20 min，回收核酸。将团块重悬于 18.5 μl DEPC 处理的 ddH_2O 中。

为防止大的团块沉淀，不要加入额外的盐。

160

17. 为将双链 DNA 产物的末端补平，加入下列试剂：

10× KFI 缓冲液	2.5 μl
（可选的）DTT（100 mmol/L）	1.0 μl
dNTP 溶液（每种的浓度为 2.5 mmol/L）	2.0 μl
T4 DNA 聚合酶	1.0 μl

在 37℃下将反应孵育 5 min，然后加入 25 μl DEPC 处理的 ddH_2O 使总体积达到 50 μl。

18. 用等体积（50 μl）的酚：氯仿抽提反应，然后向水相加入 5 μl 3 mol/L 醋酸钠或者 5 mol/L NaCl。
19. 加入 125 μl 冰冷的 95%乙醇并将溶液在－80℃孵育 30 min 来沉淀样品。4℃下在微量离心机中以最大离心速度（16 000 *g*）离心 20 min，回收核酸。

aRNA 扩增（第一轮），产量大约为 2000×

重要事项：仅能使用经过 DEPC 处理和高压灭菌的 ddH_2O；仅能使用经过高压灭菌和新打开的微量离心管、吸头或滤头。

20. 为除去会抑制 T7 RNA 聚合酶的未掺入的 dNTP，通过微量透析纯化样品。

    a. 将 DNA 团块悬浮在 20μl DEPC 处理的 H_2O 中。

b. 用 Millipore 过滤器（孔径 0.025 μm）将 10 μl 样品对 50 ml DEPC 处理的 H_2O 透析 4 h。

最好将过滤器轻轻地漂浮在装有水的灭菌的 50 ml 锥形管中。过滤器用经过“烘焙”没有 RNase 的镊子固定在试管中。为了防止材料的损失（如过滤器沉入水中），建议一次仅透析一半的样品。在透析时样品的体积会发生改变。

c. 透析完成后，将样品转移到新的、灭菌的微量离心管中。

21. 将 1/5 的透析回收样品（约 2 μl）转移到新的微量离心管中，然后加入下列试剂：

DEPC 处理的 ddH_2O	11.5 μl
10× RNA 扩增缓冲液	2.0 μl
DTT（100 mmol/L）	1.0 μl
NTP 溶液（每种的浓度为 2.5 mmol/L）	2.0 μl
RNasin（20 单位）	0.5 μl
T7 RNA 聚合酶（1000 单位/μl）	1.0 μl

将反应在 37℃温育 4 h，然后加入 30 μl DEPC 处理的 H_2O 使总体积为 50 μl。

为了分析反应产物的大小分布，按照下面信息栏中的描述建立一个“热”反应。

22. 用等体积的（50 μl）酚：氯仿抽提反应，然后加入 5 μl 3 mol/L 醋酸钠。

23. 加入 125 μl 冰冷的乙醇和 0.5 μg tRNA，并将溶液在－80℃孵育 30 min 来沉淀样品。4℃下在微量离心机中以最大离心速度（16 000 g）离心 20 min，回收核酸。

再次扩增：第一链 cDNA 的合成

24. 将步骤 23 中扩增的 RNA 的团块悬浮在 17 μl DEPC 处理的 H_2O 中。

25. 将 8～17 μl aRNA 样品在 85℃加热变性 5～7 min，然后立即在冰上冷却。

分析 RNA 产物的大小分布

1. 准备可以于分析扩增 RNA 大小分布的含有放射性标记 dNTP 的反应混合物。将约 2 μl 步骤 20 中透析回收的样品加入到一个微量离心管中并加入下列试剂：

DEPC 处理的 ddH_2O	7.5 μl
10×RNA 扩增缓冲液	2.0 μl
DTT（100 mmol/L）	1.0 μl
3 NTP（ATP、GTP、UTP 的浓度均为 2.5 mmol/L）	2.0 μl
CTP（100 μmol/L；最终浓度为 5 μmol/L）	1.0 μl
RNasin（20 单位）	0.5 μl
[α-^{32}P] CTP（3000 Ci/mmol；1 mCi/100 μl）	3.0 μl
T7 RNA 聚合酶（1000 单位/μl）	1.0 μl

将反应在 37℃下温育 3.5～4 h。

2. 向 2～4 μl 热的 aRNA 产物中加入下列试剂：

DEPC 处理的 ddH_2O	至 20 μl

37%福尔马林	3.5 μl
10× MOPS	2.0 μl
去离子甲酰胺	10 μl

3. 将样品在85℃下热变性5~7 min，然后加入2 μl 10×凝胶上样染料。
4. 用1%变性凝胶电泳分析样品，直到染料的前沿几乎到达凝胶的底部。

   不要将染料前沿迁移出凝胶，否则凝胶盒会被放射性污染。
5. 为除去未掺入的标记物，将凝胶在10%TCA中轻轻晃动1~2 h，每隔15~20 min更换10% TCA溶液。
6. 按照以下方法通过毛细作用将过夜干燥的凝胶进行印迹转移：
   a. 将凝胶放置在parafilm或Saran Wrap上面。
   b. 用两张3 M Whatman滤纸和纸巾覆盖在凝胶上面。
   c. 将重物放置在整个装置上面。
7. 第二天凝胶应该被展平和干燥。用Saran Wrap包裹凝胶并将其平放在感光成像屏或胶卷上，在−80℃放置6~12 h。关于预期的aRNA大小分布数据的示例见图3-8。

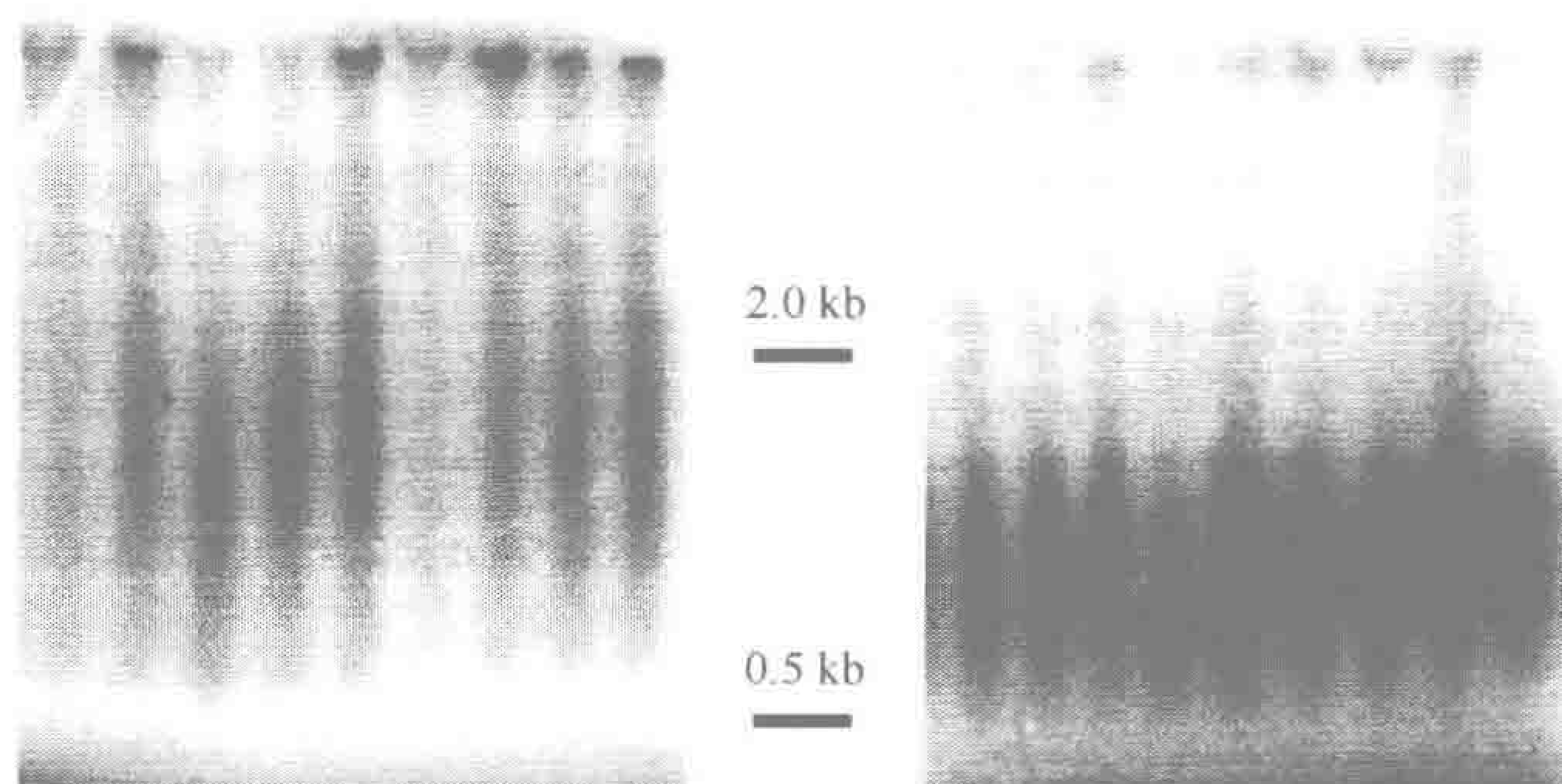


图3-8　来自一轮和二轮aRNA扩增产物的大小分布

在aRNA扩增后，其产物的长度可以在变性琼脂糖凝胶电泳上观察。（左）第一次扩增产物；（右）第二次扩增产物。aRNA长度分布表明了cDNA合成和aRNA扩增是否良好。正常情况下，mRNA的平均长度为1~2 kb。从这些图中可以看到，aRNA与父本RNA的大小相似，这说明aRNA的质量很好

26. 向变性的样品加入下列试剂：

10× RT缓冲液	3 μl
DTT（100 mmol/L）	3 μl
dNTP溶液（2.50 mmol/L）	3 μl
随机六聚体（100 ng/μl）	1 μl
RNasin（40单位/μl）	1 μl
AMV-RT（50单位）	2 μl
DEPC处理的 H_2O	加至30 μl

将反应在37℃温育1 h，然后加入20 μl DEPC处理的H_2O使总体积为50 μl。

27. 用等体积的（50 μl）酚∶氯仿抽提反应，然后向水相加入5 μl 3 mol/L醋酸钠。
28. 加入125 μl冰冷的95%乙醇，并将溶液在−80℃孵育30 min来沉淀样品。4℃下在微量离心机中以最大离心速度（16 000 *g*）离心20 min，回收核酸。将团块重悬于12 μl DEPC处理的H_2O中。

再次扩增：第二链cDNA的合成

29. 将样品在85℃加热变性5 min，然后将它放在冰上。
30. 向变性的样品中加入下列试剂，准备第二链反应：

10× 第二链缓冲液	2.0 μl
oligo（dT）-T7扩增引物（100 ng/μl）	1.0 μl
dNTP溶液（每种2.50 mmol/L）	2.0 μl
DTT（100 mmol/L）	1.0 μl
T4 DNA聚合酶（1单位/μl）	1.0 μl
Klenow片段（2单位 /μl）	1.0 μl

将反应在14℃孵育4～16h，然后加入30 μl DEPC处理的ddH_2O使总体积为50 μl。

31. 用等体积的（50 μl）酚∶氯仿抽提样品，然后加入5 μl 3 mol/L醋酸钠。
32. 加入125 μl冰冷的95%乙醇，并将溶液在−80℃孵育30 min来沉淀样品。4℃下在微量离心机中以最大离心速度（16 000 *g*）离心20 min，回收核酸。将团块重悬于18 μl DEPC处理的ddH_2O中。
33. 为补平双链DNA产物的末端，加入下列试剂：

10× KFI缓冲液	2.5 μl
dNTP溶液（每种2.50 mmol/L）	2.5 μl
T4 DNA聚合酶（1单位/μl）	1.0 μl
Klenow片段（2单位 /μl）	1.0 μl

将反应在37℃温育30 min，然后加入25 μl DEPC处理的ddH_2O使总体积为50 μl。

34. 用等体积的（50 μl）酚∶氯仿抽提样品，然后加入5 μl 3 mol/L醋酸钠。
35. 加入125 μl冰冷的95%乙醇，并将溶液在−80℃孵育30 min来沉淀样品。4℃下在微量离心机中以最大离心速度（16 000 *g*）离心20 min，回收核酸。将团块重悬于18 μl DEPC处理的ddH_2O中。

aRNA扩增（第二轮）

36. 根据aRNA扩增方法进行操作（第一轮，步骤20）

用这些方法产生的aRNA探针可以被标记并用于筛选微阵列。图3-9中显示了使用从单个细胞产生的aRNA探针获得的一个代表性结果。在这些研究中，经过两轮扩增得到的aRNA被转变为荧光标记的cDNA，并用于筛选一种Incyte GEM微阵列。这些

数据的散点图（图 3-10）表明细胞中的大多数 mRNA 的丰度位于 mRNA 在两种细胞内丰度相等的恒等线（line identity）的一个标准误差范围之内。如果有充足的探针可以和微阵列很好地杂交（荧光信号大于 400 单位），这些数据可重复产生。

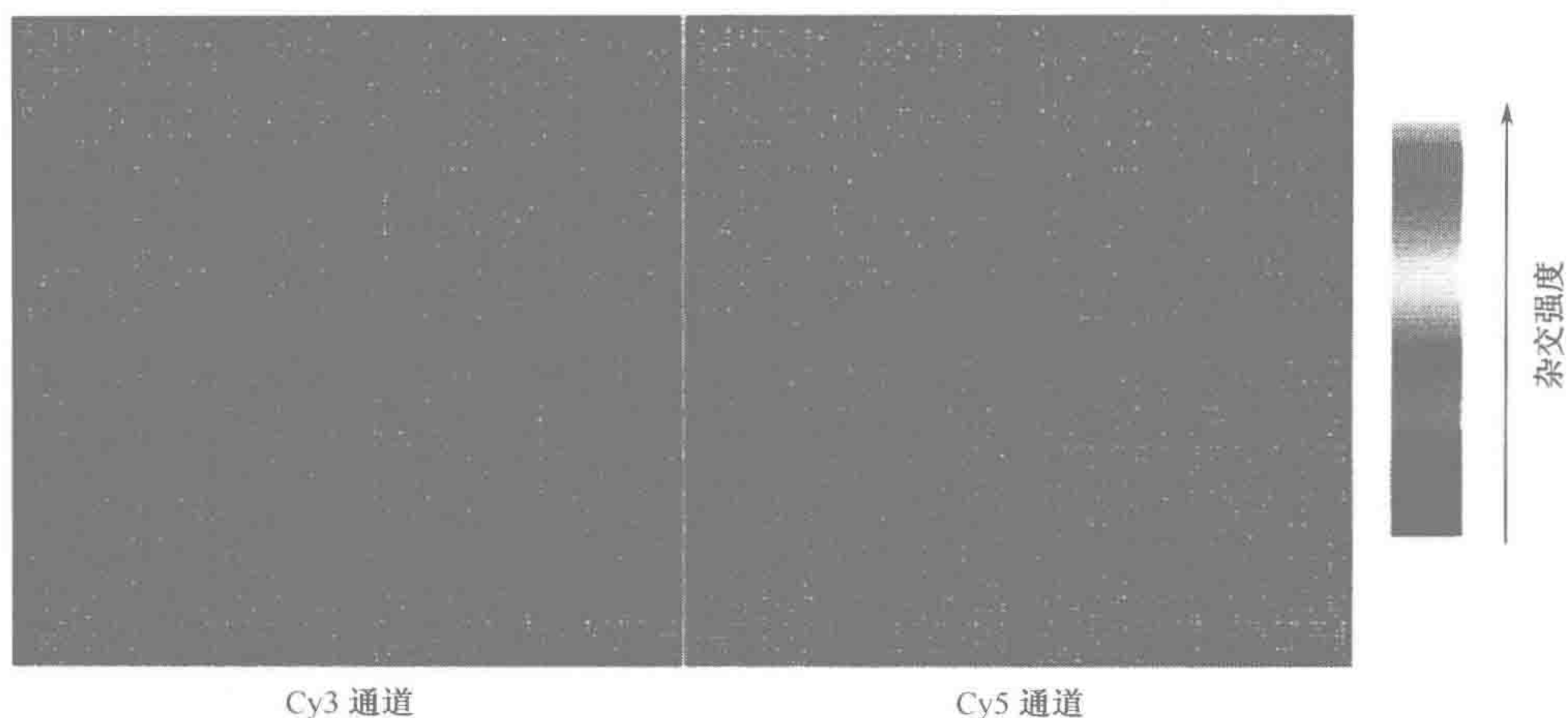


图 3-9　对来自单个细胞的 aRNA 进行筛选得到的微阵列结果

这些结果显示了一个典型的微阵列（Incyte 微阵列）与由单个细胞产生的 aRNA 制备的荧光标记 cDNA 杂交后的情形（见彩图）

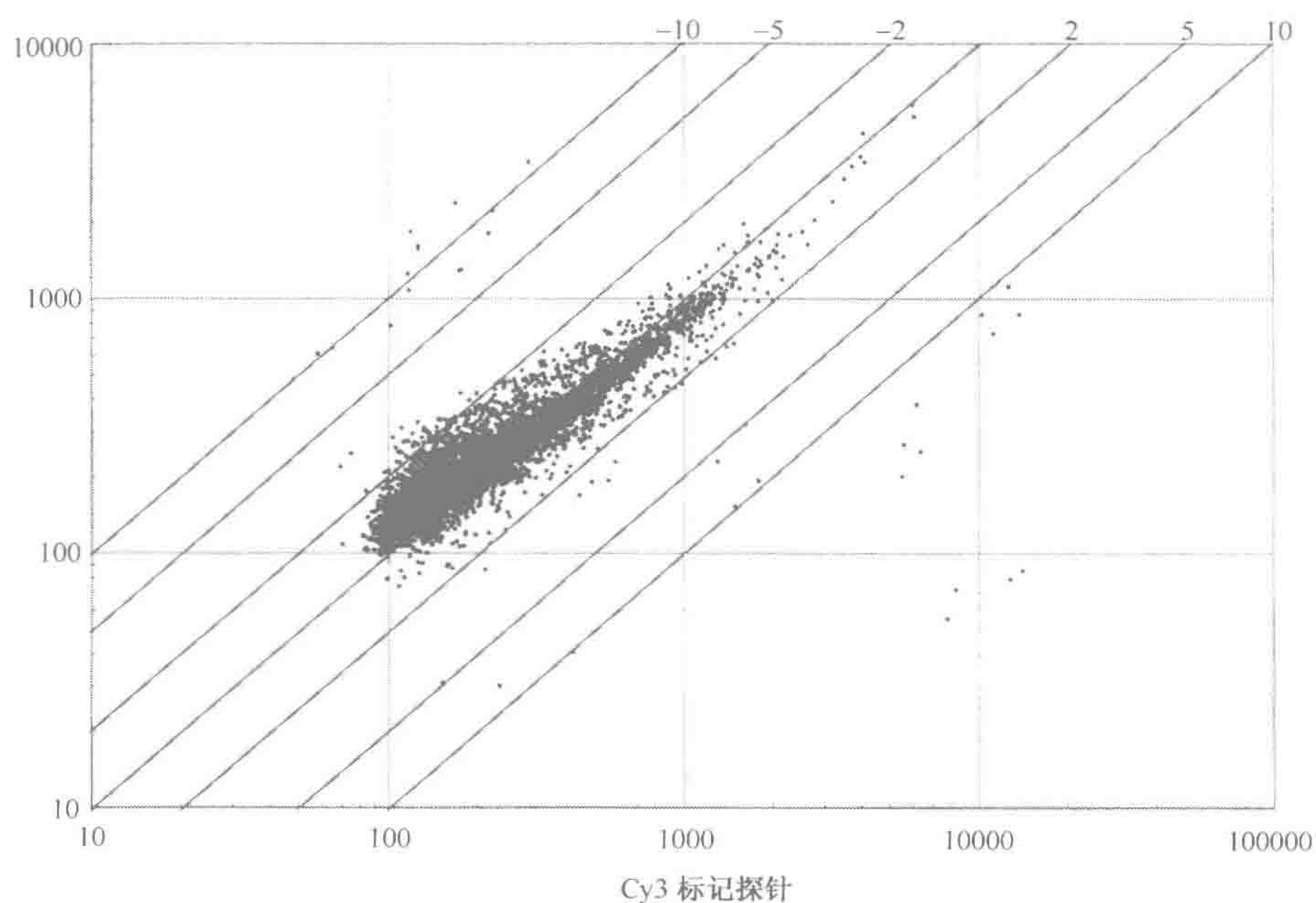


图 3-10　单个细胞的微阵列实验结果的散点图（表达结果来自 Cy5 标记探针和 Cy3 标记探针）

该散点图显示了来自两种不同类型细胞的各种 mRNA 的相对丰度。任何落在 45°对角线上的点表明那些 mRNA 的丰度在两种细胞间是相同的。这些数据的相关系数为 0.91

疑难问题解答/建议

- 扩增步骤必须严格保持没有 RNase。
- 如果酶（RNasin 和 T7）的总体积大于反应混合物总体积的 1/10，酶储存液中的甘油会抑制 T7 RNA 聚合酶。过量的 tRNA 也会抑制 T7，还有 dNTP 的竞争性抑制（透析模板的根本原因）。
- 如果反应混合物中 [α-^{32}P] CTP 的浓度升高，那么 aRNA 产物的平均长度会降低。这个问题主要是由合成后的放射分解引起的。

替代方案：组织切片中 RNA 的原位转录

在含有 oligo (dT)$_{24}$- T7 引物（主要方案中使用的）、dNTP 和反转录酶的反应混合物中，通过温育选择的组织切片可以在原位将 mRNA 直接拷贝成 cDNA。从切片回收的 cDNA 第一链产物用于制备双链 cDNA 和随后的 RNA 扩增。图 3-6 示例了整个过程。

附加材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

载体 DNA（*E. coli* DNA；10 mg/ml）

DTT（100 mmol/L）〈!〉

去离子甲酰胺〈!〉

IST 缓冲液

　　50 mmol/L Tris（pH 8.3）

　　120 mmol/L KCl

　　6 mmol/L $MgCl_2$

KCl（1.2 mol/L）

$MgCl_2$（1 mol/L）

NaOH（0.2 mol/L）/0.1% SDS〈!〉

　　可选的，见步骤 10

20× SSC

Tris（0.2 mol/L，pH 7.0 和 1 mol/L，pH 8.3）

水浴，预置为 95℃

方法

与扩增引物杂交

1. 如果组织切片是冰冻的，融化至室温并让其干燥。
2. 准备足量的下列杂交液。每 1 ml 中含：

甲酰胺	0.5 ml（50%甲酰胺）
20× SSC	0.25 ml（5× SSC）
oligo (dT) -T7 引物（100 ng/μl）	10 μl（最终浓度 1 ng/μl）
DEPC 处理的 ddH_2O	0.24 ml

3. 吸取杂交液到组织切片上（准确的体积不很关键），并在一个湿润的盒中让引物与切片在室温下杂交过夜（8～12 h）。

原位转录（IST）

4. 从切片上吸走杂交液，室温下将切片在大量的 2×SSC 中漂洗 2 h 以洗掉多余的引物。
5. 从漂洗液中取出切片，吸取 1× IST 缓冲液（无 RNase）到切片上。
6. 室温下温育切片 30 min 以改变切片上缓冲液的条件。

   这个步骤对于有效的 IST 操作非常重要。

7. 从切片上吸走 IST 缓冲液并用 IST 反应混合物代替它。

每毫升 IST 缓冲液所需储液		IST 反应混合物的组分
1 mol/L $MgCl_2$	6 μl	6 mmol/L $MgCl_2$
1 mol/L Tris（pH 8.3）	50 μl	50 mmol/L Tris（pH 8.3）
1.2 mol/L KCl	100 μl	120 mmol/L KCl
100 mmol/L DTT	71 μl	7 mmol/L DTT
DEPC 处理的 H_2O	660 μl	
dNTP 溶液（2.5 mmol/L）	100 μl	250 μmol/L 四种 dNTP
RNasin（40 单位/μl）	3 μl	0.12 单位/μl RNasin
AMV-RT（20 单位/μl）	10 μl	1 μl RT/100 μl 混合物

重要事项：将 AMV-RT 单独加入每支试管；加入 RNasin 之前，将 DTT 加入到反应混合物之中以形成合适的还原性环境；最后加入反转录酶；将 IST 反应在 37℃下温育 90 min。

8. 用大量的 0.5×SSC（不需要是无 RNase 的）在室温下清洗切片 2 h，或者在 4℃过夜，然后让切片在空气中干燥。

   这些载玻片现在可以用于继续处理，或者也可以在室温储存数周，或在 4℃长期保存。将载玻片在载玻片盒中保存。

从组织切片中回收 cDNA

9. 通过使用微量移液器刮取并转移含有原位转录的 cDNA 的固定细胞，将细胞吸入到含有 10 μl DEPC 处理的 H_2O 的离心管中。

   作为替代，除去周围的组织区域，将细胞留在载玻片上，然后用微电极将细胞转移到微量离心管中。

10. 95℃加热 5 min 从细胞碎片中分离 cDNA，然后迅速冷却。

    作为替代，加入 10 μl 0.1 mol/L NaOH 或 0.1% SDS 使 cDNA 化学变性。

11. 加入 10 μl 0.2 mol/L Tris（pH 7.0）和 1 μl 10 mg/ml *E. coli* DNA（载体）到 cDNA 中，并使用乙醇沉淀回收 cDNA。

    不要使用酚：氯仿抽提，因为残留的蛋白质将充当载体并有助于沉淀。

12. 除去乙醇，立即倒转试管，将团块在空气中干燥 10 min。将团块重悬于 20 μl DEPC 处理的 ddH_2O 中。
13. 将 cDNA 团块溶解在 20 μl DEPC 处理的 H_2O 中并前进到主要方案的步骤 12 继续 cDNA 第二链的合成。

附加方案：切片的吖啶橙染色

在对从固定组织切片上回收的单个细胞中抽提的 RNA 进行分析之前，用一张切片去检查一下经过组织的收集和固定之后 RNA 是否保存下来是非常必要的。通过将带有合适切片的载玻片浸入吖啶橙溶液，这很容易做到（图 3-11）。在适当的 pH 条件下，吖啶橙将单链核酸（如 RNA）染成橘黄色，而双链 DNA 染成绿色。如果在荧光照射方式下检测到染色的组织切片中有橘黄色出现，那么就存在 RNA。大多数染色来自于核糖体和其他丰富的 RNA。无论如何，只要这些 RNA 种类存在，就有很好的机会。用柠檬酸-磷酸钠缓冲液处理过的切片适合于使用原位转录单个细胞的 aRNA 的方案。

附加材料

吖啶橙

用柠檬酸-磷酸钠缓冲液配制吖啶橙溶液（0.6 μg/ml）。

柠檬酸-磷酸钠缓冲液

60 mmol/L 柠檬酸

80 mmol/L Na_2HPO_4（pH 4.1）

方法

1. 用 DEPC 处理的 ddH_2O 漂洗组织切片 5 min。
2. 将切片在柠檬酸-磷酸钠缓冲液中孵育 5 min。
3. 将切片在吖啶橙中孵育 15 min。
4. 在柠檬酸-磷酸钠缓冲液中漂洗切片 3 次，每次 5 min。
5. 用荧光显微镜在 488 nm 的波长下观察单链 RNA（橘黄色）和双链 DNA（绿色）。

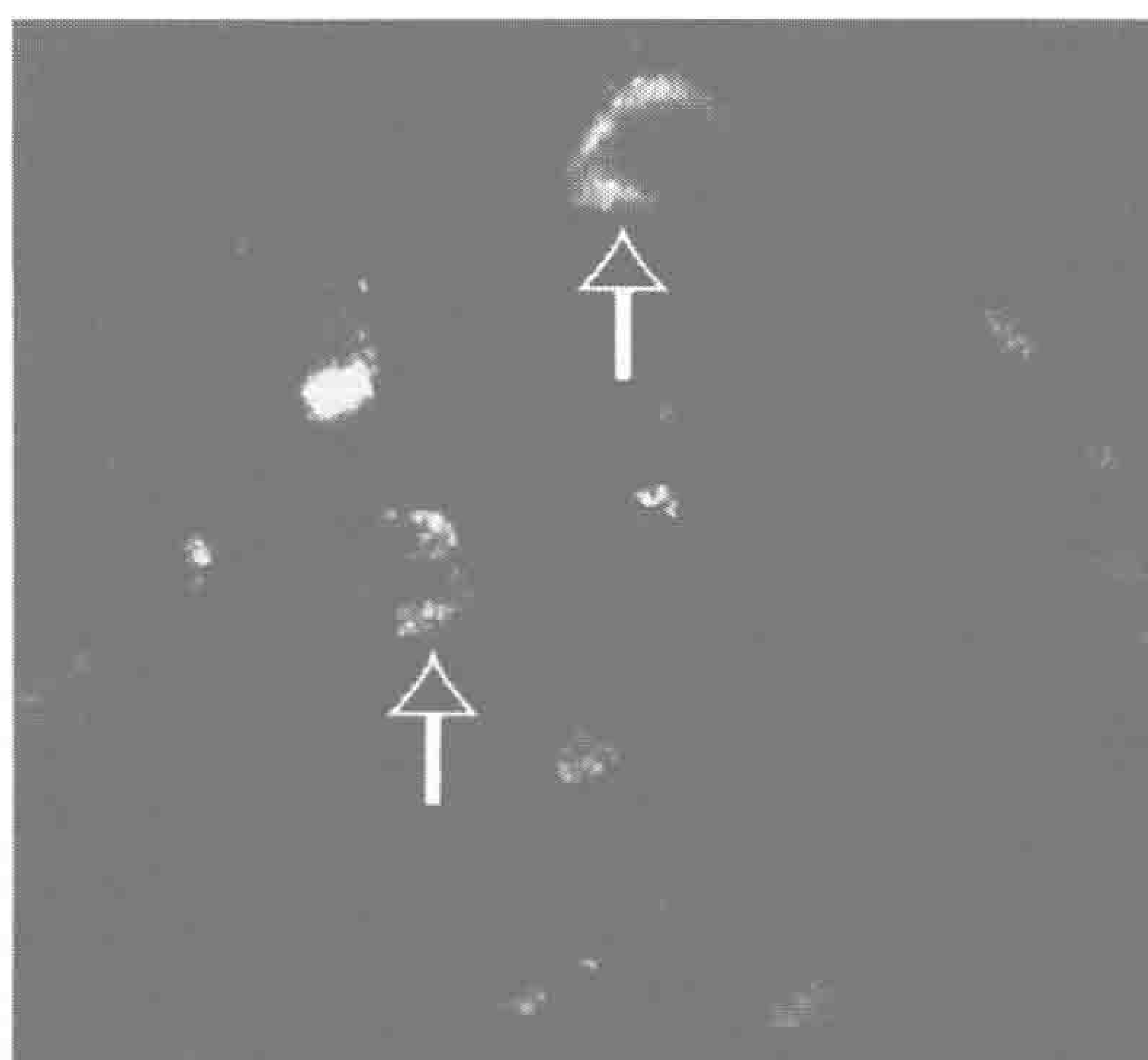


图 3-11　切片的吖啶橙染色用于观察 mRNA

在这张切片中可见到吖啶橙染色的 mRNA。这种方法应用于存档的组织标本（见彩图）

方案 12

RNA 通用参考集

Jonathan R. Pollack（病理学系，斯坦福大学医学院，斯坦福，加利福尼亚 94305）

使用微阵列进行基因表达测量时，对两种不同 mRNA 制备物进行标记，使用不同的荧光基团，并且同时和一个 cDNA 微阵列杂交。对于阵列中每一个基因特异性的单元，两种荧光强度的比值反映了在两种制备物中特定 mRNA 的相对丰度。

使用双色荧光标记的重要特性是能够有力地消除点样到微阵列上 DNA 数量的变化、杂交严格性（盐浓度、温度和杂交时间）的变化和杂交过程中标记样品局部浓度的变化带来的影响。这些参数的改变都同等地影响荧光比值的分子和分母，因此不会影响测量的荧光比值。

在包括多种不同 RNA 制备物的研究中，将所有成对组合的样品进行杂交是不切合实际的。解决的方法是将每种实验样品和相同的通用参考 RNA 制备物进行杂交，并计算出每种 mRNA 相对于参考的水平。然后根据比值，间接计算出不同实验样品中 mRNA的相对丰度。

对于许多实验，有一个“天然的”RNA 参考样品可供使用，例如，当定量 mRNA 的丰度随时间变化时（参比＝零时刻的样品）或者在药物处理后（参比＝未接受药物处理的对照）。但是，在特定的情况下，如测量不同肿瘤样品中的 mRNA 丰度时，没有天然的参考存在。在这些情况下，可以使用一个 mRNA“通用参考集”（Eisen and Brown 1999）。如果每个测试（如肿瘤）样品和相同的通用参考集进行对比，那么这些样品中基因表达的相对强度就可以计算出来。在这种类型的实验中，通用参考集成为所有测量的荧光比值的分母。

一个理想的通用参考集应该和微阵列中的全部或者几乎全部 DNA 单元以足够的信号强度进行杂交，应该经济且易于制备，并且容易复制以便随时间变化和在不同位置之间进行样品的比较。要注意的是，通用参考集不需要有任何特定的生物学目的或者特性，并且不需要包含来自所研究的每种细胞类型的 RNA。但是，一个好的参考集应该从阵列上大多数单元产生可读的信号。不仅如此，为了使参考集适合于大多数研究人员，还应该用容易得到的、经过证实的来源制备。实际上，包含从培养细胞的稳定细胞系合并的 RNA 的通用参考集能够满足这些标准。

本方案概述了一个广泛用于斯坦福大学 cDNA 微阵列研究单位的 RNA 通用参考集的制备（Perou et al. 1999，2000；Ross et al. 2000）。该参考集从 11 个不同的稳定细胞系制备（表 3-6），这些细胞系大多数来自于美国菌种保藏中心（ATCC）。这些细胞系容易生长，并且结合在一起可以和杂交到 cDNA 微阵列上大多数单元杂交。尽管用于参考集的最佳细胞系的数目还未通过实验确定，可能还是不要使用太多的不同细胞类型较好。合并的 RNA 可以稀释到那些只在集合中的一种或者一些细胞系中表达的低丰度转录物不可见的程度。

大多数具备 DNA 微阵列基础设施的实验室应该能够复制斯坦福的通用参考集。通常，制备一批通用参考集一般需要 2～3 个月，生长 11 个细胞系和纯化足量的 RNA 用于几千个微阵列杂交。但是，对于一些研究人员，一个商品化的参考集可能是最好的选择。Stratagene Inc. 公司销售从十种人类细胞系中抽提的 RNA 通用参考集（Novoradovskaya et al. 2000）。但是，与斯坦福通用参考集不同，几乎没有关于细胞系和他们来源的详细说明。

使用斯坦福和 Stratagene 的通用参考集都能够在实验室内部或是实验室之间对微阵列结果进行合理的比较。其他可能的参考集，可以满足同样的用途，包括由合并的 cDNA体外转录产生标记的基因组 DNA 或 RNA，但不常用。

为何使用通用参考集？

为了确定在肿瘤和正常细胞中表达水平不同的基因，有两种可能的实验设计：

- 肿瘤 vs. 肿瘤 vs. 正常

  样品：T1，T2，N1，N2

  可供选择的对比：

  T1 vs. N1，T2 vs. N2，T1 vs. N2，T2 vs. N1，T1 vs. T2，N1 vs. N2

  该矩阵提供了所有可能组合的对比，并且优势在于所有的问题都可通过直接对比处理：一个肿瘤和另一个相比如何？肿瘤组织和正常的相比如何？正常的和正常的相比如何？但是，这种方法只有当样品的数量非常少的时候才是实用的，因为组合的数量随着样品数目成指数增长。

- 肿瘤 vs. 参考，正常 vs. 参考

  样品：T1，T2，N1，N2，R

  可供选择的对比：

  T1 vs. R，T2 vs. R，N1 vs. R，N2 vs. R

  这里，全部的样品相对于一个通用参考 RNA 进行的测试和比较是通过相

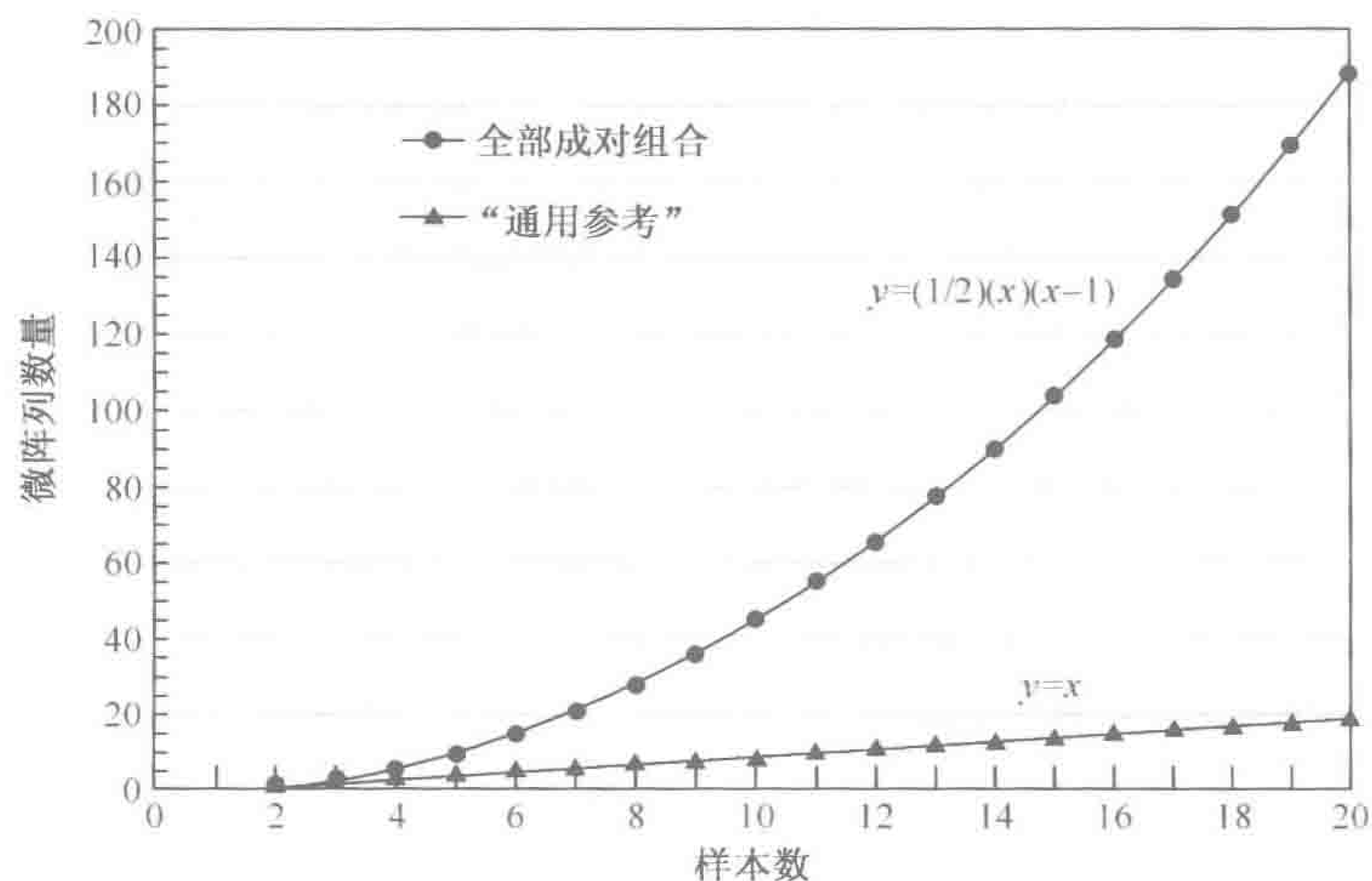


图 3-12 进行全部成对组合 vs. 使用一个“通用参考”的不同之处

170

对于参考间接进行的。组合的数目易于处理，因为它们随着样品的数目线性增长。此外，比较可以在使用了通用参考集的实验或研究人员之间进行。缺点是推论是在样品之间通过参考间接得到的，潜在地削弱了实验解决问题的能力。因此，理想的参考集应该是大多数人易于得到、获取数量不受限制的，并且对于范围广泛的基因在微阵列中都能提供信号的。这些问题将在第 7 章中更加详细地讨论。

图 3-12 说明了这两种实验设计中的差异。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度并过滤灭菌。

ATRA（全反式维甲酸；R2625，Sigma）

配制在 DMSO 中的 10 mmol/L 储存液。

细胞培养基

RPMI 1640（GIBCO 或与之相当的）

10% 胎牛血清（HYCLONE 或与之相当的）

青霉素和链霉素（1×，GIBCO）

磷酸盐缓冲液（PBS），不含 Ca^{2+} 和 Mg^{2+}

1×胰蛋白酶-EDTA

凝胶

琼脂糖凝胶（1%）含有 2.2 mol/L 甲醛〈!〉

详细情况，见 Sambrook and Russell (2001)。

专用设备

FastTrack 2.0 mRNA 分离试剂盒（Invitrogen）

含有 5% CO_2 潮湿气体的培养箱

传代和收获细胞的层流通风橱

细胞培养转瓶和合适的培养箱（替代组织培养皿）

组织培养皿（15 cm）

UV 分光光度计

细胞和组织

细胞系

关于斯坦福的 RNA 通用参考集中使用的细胞系的清单和说明见表 3-6。11 个细胞系中，当

前有 9 个能够从 ATCC（www.atcc.org）获取。但是，因为 NB4 和 UACC-62 当前不能由 ATCC 供应，所以从公开的来源无法复制斯坦福通用参考集。因此我们推荐用 HL-60 细胞（ATCC CCL-240）来替代 NB4 细胞，用 SK-MEL-2 黑色素瘤细胞系（ATCC HTB-68）来替代 UACC-62 细胞。由于目的是要尽可能地保留参考集的“公共性”，不推荐个别实验室中使用的其他替代。

表 3-6　斯坦福“RNA 通用参考集”中使用的细胞系

名　称	说　明	生长特点	ATCC 分类号或参考文献
MCF7	乳腺癌衍生的细胞系	黏附	ATSS HTB-22
Hs578T	乳腺癌衍生的细胞系（基质样细胞，stromal-like）	黏附	ATSS HTB-126
NTERA2	畸胎瘤衍生的细胞系	黏附	ATCC CRL-1973
Colo205	结肠癌衍生的细胞系	混合	ATCC CCL-222
OVCAR-3	卵巢癌衍生的细胞系	黏附	ATCC HTB-161
UACC-62	黑色素瘤衍生的细胞系	黏附	Stinson et al.（1992）
MOLT-4	T 细胞白血病衍生的细胞系	悬浮	ATCC CRL-1582
RPMI 8226	多发性骨髓瘤衍生的细胞系	悬浮	ATCC CCL-155
NB4＋ATRA	急性早幼粒细胞白血病衍生的细胞系	悬浮	Lanotte et al.（1991）
SW872	脂肪肉瘤衍生的细胞系	黏附	ATCC HTB-92
HepG2	肝癌衍生的细胞系	黏附	ATCC HB-8065

方法

细胞系生长

表 3-6 中列出的 11 个细胞系在单层或悬浮培养中都易于保持和生长。本方案是为生长在培养皿里的细胞而设计的。如果悬浮培养物是在长颈瓶或转瓶中生长，要相应地调整方案。

1. 37℃下，在含有 5%CO_2气体的培养箱中，用 15 cm 的组织培养皿盘，将 11 个细胞系保持在添加 10%胎牛血清的 RPMI 1640 培养基中。

   当接近单层汇合时，细胞应该使用黏附细胞系的标准胰蛋白酶消化作用方法，按照约 1∶5稀释进行传代培养。

   在记录本或数据库中记录下细胞传代的次数。将早期传代的原代培养物在含有 10% DMSO的培养基中以小份保存于液氮中。

   弃去传代 10 次以上的细胞系，并用冰冻的母培养物代替它们。

   如果可能，使用同一批次的胎牛血清来培养 11 个细胞系。在参考 RNA 中，血清可能是批次间变化的一个重要的因素。记录下血清的批号和每个细胞系每次传代时使用的培养基。

2. 扩大细胞的数量，使得每个细胞系可以有 10～20 个 15 cm 组织培养皿的培养物。

3. 于收获细胞分离 mRNA 之前 48 h，更换培养基。

   培养基变化改变了基因表达的模式和强度。为了减少通用参考 RNA 制备物之间的变化，建立标准的方案用于细胞的生长和再培养（refeeding），并严格遵守。

172

4. 在收获细胞之前约 48 h，向 NB4 细胞（或替代的 HL-60 细胞）的培养基中加入全反式维甲酸（ATRA）至终浓度为 10 μmol/L。

   NB4 和 HL-60 来自人类的早幼粒细胞白血病例。加入 ATRA 诱导，在 48 h 后，粒细胞分化的特异性基因表达同时伴随着细胞的成熟。

5. 在约 80%的培养细胞单层汇合时，收获细胞用于分离 mRNA。

   当细胞培养物接近单层汇合时，基因表达的模式发生变化。为了减少通用参考 RNA 制备物之间的变化，建立标准的方案用于细胞的生长和收获，并严格遵守。

制备 mRNA

6. 使用 mRNA 分离试剂盒，如 FastTack 2.0（Invitrogen）从细胞系中制备 mRNA。
   a. 从培养物中吸走培养基。
   b. 使用生产商提供的裂解缓冲液来裂解培养皿表面的细胞。
   c. 使用 oligo（dT）-纤维素批量制备色谱法分离 mRNA。

   许多方法可以用于从细胞系中制备总 RNA 或者 mRNA。斯坦福研究组一般使用 FastTack 2.0 (Invitrogen) mRNA 分离试剂盒。FastTack 试剂盒可产生高质量的 mRNA 用于标记和与 DNA 微阵列杂交，价格相对便宜的替代品也可能同样有效的。

   这里给出的是方法的概述，详细情况见方案 6。

7. 用 UV 分光光度计或者 RiboGreen 染色来确定 mRNA 的产量和纯度（见 **RNA 的定量**信息栏）。
8. 通过甲醛-琼脂糖凝胶上的变性凝胶电泳来评价 mRNA 的质量和完整性（见**哺乳动物、植物和细菌 RNA** 信息栏）。
9. 如果分离的 mRNA 的数量和质量均是令人满意的，合并等量（质量）的来自每个细胞系的 mRNA 并混合均匀。将合并的 mRNA 混合物分装成小的等份（10～20 μg），将它们储存在－80℃或按照 **RNA 的保存和回收**信息栏中所述保存。

校正不同批次的通用参考 RNA

尽管我们努力使细胞系的培养和 mRNA 的分离按照一致和标准的方式进行，但是细胞处理或者 mRNA 定量的细微差别，也可能改变参考集中不同 mRNA 的丰度。为了将使用不同批次参考集的实验结果关联起来，可能需要进行定量和校正批次间的差异。校正可以通过用 Cy3 和 Cy5 标记不同批次的参考集，并将它们按照成对交叉的方式（pair-wise cross）与 DNA 微阵列杂交。在这个实验重复数次以后，从阵列上的每一个基因能够计算出精确的批次特异性的"校正"系数。这些系数可以用来校对由于通用参考集中批次之间的变化而导致的系统偏差。

173

PART Ⅱ 标记和杂交

导言

在 20 世纪 70 年代中期进行了首次基因表达的整体性（global）测量。从培养哺乳动物细胞的细胞系分离的多核糖体 RNA 在体外转录为放射性标记的单链 cDNA，然后

与它的大量未标记的 mRNA 模板进行杂交。杂交动力学表明，mRNA 包含了三种动力学类型，且它们在 mRNA 总量中所占丰度各不相同（Biship et al. 1974；Getz et al. 1975，1977）。含量最为丰富的转录物占 mRNA 质量的 30%～90%，但是序列复杂度只占到 0.1%。中等丰度的 mRNA 约占 mRNA 质量的 50%～60%和序列复杂度的 2%～5%。稀有 mRNA 所占的质量几乎可以忽略，但却提供了大于 85%的多核糖体 RNA 的序列复杂度（表 3-7）。

表 3-7 哺乳动物和植物细胞中 mRNA 的丰度种类

mRNA 种类的总数	$1.0\times10^4\sim3.4\times10^4$
序列复杂度	$1\times10^7\sim3\times10^7$ 核苷酸
每个细胞中 mRNA 分子	
高丰度的 mRNA	
种类的数量	7～10
序列复杂度	约 $1\times10^4\sim3\times10^4$ 核苷酸
分子的数量/细胞	4500～10 000 分子/细胞
中等丰度的 mRNA	
种类的数量	500～800
序列复杂度	$2\times10^5\sim1.5\times10^6$ 核苷酸
分子的数量/细胞	300～1000 分子/细胞
低丰度的 mRNA	
种类的数量	10 000～34 000
序列复杂度	$1\times10^7\sim3\times10^7$ 核苷酸
分子的数量/细胞	3～100 分子/细胞

数据来自 Bishop 等（1974）；Geta 等（1975）；Goldberg 等（1978）。

据估计，约 85%表达的基因在每个细胞中存在的拷贝小于或等于 5 个（<1：100 000）；细胞中约 97%质量的 mRNA 仅仅包含约 20%转录的基因。

此后 25 年的工作充分证实了这些一般性的结论。在四分之一个世纪的时间里，通过综合 Northern 印迹法、RNase 保护和近来的实时 PCR，隶属于每种丰度的大量 mRNA被分类、作图和定量。然而，很明确的是，细胞内转录整合的全面图谱不可能根据单个 RNA 的研究自下向上来建立，因此需要有能够同时检测许多基因表达的新技术。在 20 世纪 90 年代早期，差异显示（differential display）在部分上似乎是有潜力的解决这一问题的，但是，在最近十年的中期，大约是和第一篇关于 DNA 微阵列的文章发表在同一时间，差异显示方案由于缺乏效力和没有研究出稳健的自动化方法而告失败。

DNA 微阵列是当前在 RNA 谱分析中占优势的技术，已经符合关于稳健性和效能的全部合理测试，并且对于小的生物已经实现了定量单个细胞内转录产物和描述对于特定的外部刺激产生相似响应方式的各类基因的承诺。

174

但是这些仍然是不成熟的技术，并且在得到真核生物中基因表达全面和可靠的数据库之前，还需要若干年的工作。目前，我们关于哺乳动物和植物的复杂组织和细胞中整体性基因表达的描绘是不完整和低分辨率的：前者是由于只有细胞内转录单位的一个子集表示在现在使用的 DNA 微阵列上；后者是由于隶属于最低丰度类型的转录物，处于或者刚好在标准的微阵列技术的灵敏度极限之下。面对面的直接比较表明，当前 DNA 微阵列的灵敏度几乎等同于寡核苷酸芯片、Northern 印迹法和其他 RNA 分析方法。（Chen et al. 1998；Bertucci et al. 1999；Wang et al. 1999；Yang et al. 1999；Taniguchi

et al. 2001）。

一个有用的微阵列实验灵敏度的衡量指标是通过与同源的、点样到阵列单元上的DNA杂交进行可靠检测的最微量的mRNA（smallest fraction of mRNA）（Bertucci et al. 1999）。在理想的实验中，这个微量是百万分之一到百万分之二，足以检测一种在每个细胞中存在的量为几个分子的RNA。mRNA与典型的DNA微阵列上显示单元（display element）杂交后发出的信号强度受到很多变量的影响，包括下面所描述的内容。

RNA制备物的质量

在纯化的各个阶段，采取保护措施避免受到RNase的污染非常重要。被RNase部分降解的mRNA可能会产生比和它们的同源的显示单元要小的、截短的cDNA，由此导致信号强度的减小。阵列单元与成对的Cy3和Cy5标记的靶分子杂交所产生的光谱信号的差异反映两种RNA制备物质量的不同，而非转录的调节。使用oligo（dT）作为引物从总RNA或者纯化的poly（A）$^+$ mRNA合成荧光标记cDNA可以减小这个问题。除非RNA制备物严重降解，长度约为300～500 bp的靶分子，从转录物的3′端产生，不太可能是截短的。据报道，短于300 bp的靶分子在微阵列实验中产生不可靠的数据（Yang et al. 1999）。RNA制备物的质量可以根据**哺乳动物、植物和细菌RNA**信息栏中描述的方法进行评价。

引物属性

一些或许是大多数的研究人员更喜欢使用总RNA作为起始材料进行直接荧光标记。在这种情况下，使用oligo（dT）作为引物进行第一链的合成是必需的，并且荧光标记的cDNA产物偏向于mRNA 3′区域的序列。作为替代，单独使用oligo（dT）作为引物，或者共同使用oligo（dT）引物和随机引物的反应中，纯化的poly（A）$^+$或者多核糖体mRNA能够用作直接标记的模板。随机引物增加了荧光标记cDNA的产量和序列的复杂度，不需要选择进行mRNA序列的反转录，而且偏向于mRNA 3′区域的序列。

175

然而，经常污染mRNA制备物的其他类型的RNA（核糖体和低分子质量RNA）也会被拷贝到标记的cDNA，这将增加杂交的背景。当阵列单元是由全部的cDNA序列或者含有预测可读框（ORF）的基因组DNA的扩增片段组成时，随机引物是有利的。当阵列单元由来自mRNA 3′区域衍生的序列组成时，必须使用oligo（dT）引物或者锚定引物。

靶分子上携带的荧光基团的类型、密度和数量

有许多种方法可以用来将荧光生色团掺入或者连接到靶序列上。一种常用的方法是直接标记，即在反转录酶催化的反应中结合羰花菁染料Cy3和Cy5的核苷酸（Yu et al. 1994；Zhu et al. 1994；Zhu and Waggonor 1997）直接掺入cDNA第一链，或者在*E. coli* DNA聚合酶的Klenow片段催化反应中掺入cDNA的第二链。在两种情况下，目标都是要使标记密度达到平均8个碱基中有一个碱基携带了荧光标记。更高的标记密度会起反作用，因为相邻荧光核苷酸分子之间的猝灭减少了荧光的产量和信号强度

(Randolph and Waggnner 1997)。

广泛使用的 Cy3 和 Cy5 标记的 dNTP 作为直接标记 cDNA 的试剂可能不会持续更久。许多替代的染料（Alexa 荧光染料）刚刚问世（Molecular Probes，Inc)，它们有更强的荧光，改良的光稳定性，并且更能耐受 pH 的变化（Panchuk Voloshina et al. 1999)。在面对面的直接比较中，用 Alexa 456 标记的 cDNA 产生的信号要比 Cy3 标记的 cDNA 产生的信号高出约 2.5～3 倍（Wildsmith et al. 2001)。关于进一步的信息，见 **ALEXA DYES** 信息栏。

RNA 模板反转录使用的酶

热稳定的、DNA 依赖的 DNA 聚合酶将结合了青色素的 dUTP 掺入能力改变 2～3 倍（Panchuk Voloshina et al. 1999)。反转录酶的生产商也作出类似的声明。现在最常用于合成荧光标记的 cDNA 靶分子的酶是 SURERSCIPT Ⅱ RT(GIBCO)，它是 Moloney鼠白血病病毒（Mo-MLV）反转录酶经过基因工程改造后的形式，没有 RNase H 活性，而且不受 rRNA 或 tRNA 的明显抑制。然而，由于酶分子受到大的青色素加合物的空间阻遏，标记反应必须按照生产商的建议孵育 2h 以上。

尽管 SURERSCIPT Ⅱ RT 已经成为直接标记 cDNA 的标准酶，但许多其他形式的反转录酶也可以购买到，并且根据生产商提供的资料判断，可能与 SURERSCIPT Ⅱ RT 效果等同或者要好于它（如来自 Amersham Pharmacia Biotech 的 CyScribe 和 QIAGEN的 Omniscript)。

阵列单元中 DNA 的密度和数量

在杂交中，每种固定的 DNA 探针必须超过与其互补的标记靶序列的十倍或更多。这种安排确保了在很宽的靶序列浓度范围内，杂交将会遵循拟一级反应动力学进行。任何一个单元发出的信号强度取决于杂交混合物中同源靶序列初始的浓度。因此，靶序列浓度增加两倍将会使产生的信号增加两倍，并且较丰富的 RNA 将会比低丰度 RNA 产生更强的信号。当固定的探针 DNA 的量成为限制因素时，系统的动态范围变窄，而且对表达差异的估计也缩小（Heller et al. 1997；Yue et al. 2001)。表达分析中探针(cDNA或 PCR 产物）印刷到微阵列上的浓度通常是 100～500 ng/μl(Schena et al. 1996；Welfhrd et al. 1998；Deyholos and Galbralth 2001)。使用中，微阵列系统的动态范围可能更多的是由扫描仪，而不是靶 RNA 的浓度来确定的。扫描仪通常是 16 比特的机器，也就是说 2^{16}，动态范围约等于 65 000。

176

靶序列浓度

如果杂交混合物中靶序列的浓度太低，复性将会很慢，而且减弱的信号可能无法被荧光扫描仪检测到。杂交需要的最少量的靶序列应该是在样品中包含了刚好足量的最低丰度 RNA 能够在杂交后产生可读出的信号。这个数量可以用微量检测限（见 PART Ⅱ）与杂交混合物中靶分子的总数相乘来进行估计。由于随着 RNA 量的增加，信噪比和动态范围都会改善，杂交混合物中通常包含 3 或 4 倍计算的靶序列的最少数量。然而，如果靶序列的浓度太高，如此计算的最少数量大 10 倍，信号的动态范围可能会减

小，并且信噪比可能降低或者没有改善（Yue et al. 2001）。作为一个经验，在一个标准的微阵列杂交混合物中，使用的标记靶序列的量应该等同于至少 1 μg 的 poly(A)$^+$ RNA（约 100 μg 总 RNA）（见方案 20）。如果没有获得这么多的数量，在标记之前应该将 RNA 制备物扩增（Guatelli et al. 1990；Van Gelder et al. 1990；Eberwine et al. 1992；Luo et al. 1999）。关于总 RNA 线性扩增的详细情况，见方案 15 和 16。

信噪比

当信噪比大于 1.5～2.0 时，微阵列上典型的斑点样式可以被检测到。许多因素会影响信噪比，包括：

- **mRNA 制备物的质量**。因组织来源和制备方法的不同，RNA 的质量会发生变化。从质量降低的 RNA 合成的标记靶序列跟与它同源的 DNA 微阵列上单元的杂交可能是无效的，同时可能与非同源的阵列单元发生非特异性的结合。此外，通过与基底或者固定在阵列上 DNA 之间结合的分子，以及同时结合的荧光生色团，非特异性的噪声可能会带入分析中。
- **直接标记中使用的反转录酶的特性**。一些类型的反转录酶不能很好地处理富含鸟嘌呤和胞嘧啶的模板，而其他的反转录酶则不能催化携带青色素染料加合物的掺入。

177

- **靶序列上携带的荧光加合物的数量和密度**
- **标记反应中产生的错误延伸产物的比例和数量**。RNA 样品中的小片段 RNA 和 DNA 可能充当异常的引物，从而产生一类可以和许多阵列单元杂交的标记片段。
- **包含阵列单元的载玻片的表面特征和点样液的成分**。包被了聚赖氨酸的载玻片与 DNA 产生非共价结合，因此高温下在高盐度溶液中孵育时，DNA 容易流失。点样液会显著地影响 DNA 结合到载玻片表面的效率。载玻片表面的不均匀性会导致 DNA 在载玻片表面不同部位的非均匀分布。
- **杂交条件的严格性和包含在杂交反应中的封闭试剂的有效性**。去污剂和封闭试剂可以抑制非特异性的杂交，但是它们也会增加 DNA 从载玻片表面流失的速率。
- **洗涤过程的效率**。未能除去杂交溶液——特别是那些含有 SDS 的杂交液，会导致温度的上升。

关于导致信噪比变差的问题和其他原因，在本指南文中许多部分进行了讨论。这里的要点是，DNA 微阵列实验是一个统一体，由许多部分紧密结合而成，因而几乎所有阶段中的差错和判断失误，后来都会困扰研究人员。

178

方案 13

用反转录酶进行 cDNA 第一链的荧光标记

直接合成荧光标记的 cDNA 第一链是在包含 mRNA 作为模板，结合了染料的脱氧

核苷酸三磷酸盐为底物，oligo(dT)$_{12\text{-}18}$或者锚定的oligo(dT)$_{12\text{-}18}$作为引物，由反转录酶催化的反应中进行的。研究人员要在不同品牌的商品化反转录酶和连接了羰花菁或Alexa染料的核苷酸之中进行选择（见 **ALEXA DYES** 信息栏）。本方案中使用的标准组合（SUPERSCRIPT Ⅱ和 Cy3/Cy5 标记的 dUTP）可靠性很高，并且能够满足掺入高水平的荧光标记核苷酸的需求，发出的荧光有好的光谱分离，以及吸收峰与可用的激光器的发射波长匹配。

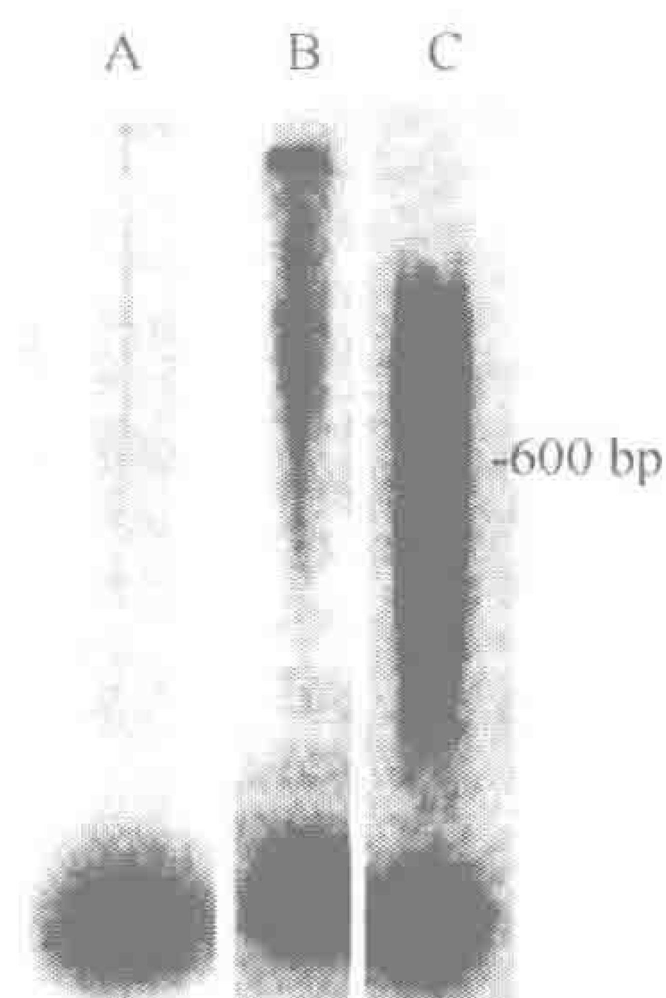


图 3-13 Cy5 标记的 cDNA 的电泳谱图

标记的 cDNA 来自所示的三种不同标记物。每种标记物的 10% 的样品在 2% 琼脂糖凝胶（Tris-乙酸盐，EDTA 缓冲液）上电泳，然后用配备了 630 nm 和 650 nm 激发光的荧光扫描仪对凝胶成像。泳道 A 和 B 显示了两个标记物的结果。泳道 A 是用片段化的 RNA 作为模板时得到的典型结果。泳道 B 显示了标记 cDNA的聚集，这主要发生于模板 RNA 只经过相分离纯化的情形。泳道 C 显示了一次成功的试验：全体荧光标记的 cDNA 显示了合理的链长范围，并且没有发生聚集

通常，在反转录酶催化的反应中，标记了染料的核苷酸掺入 cDNA 的效率要比未修饰的 dNTP 低。因为掺入青色素标记的核苷酸常常导致在邻近的一个核苷酸处发生链的终止（Zhu and Waggoner 1997），合成的荧光 cDNA 长度的中间值与反应混合物中标记了染料的核苷酸浓度成反比。在本方案描述的标记条件下，荧光 cDNA 相对较短（400～1000 个核苷酸）。

Cy5 取代的核苷酸掺入效率要比 Cy3 取代的核苷酸低。因此，建议通过 2% 琼脂糖凝胶电泳来检测 Cy5 标记的 cDNA 的链长，以此作为质量控制步骤。电泳后，用荧光扫描仪扫描凝胶。图 3-13 中显示了两个失败标记和一个成功标记的电泳谱图。

标记反应的效率和标记产物的质量可以通过将荧光标记的 cDNA 和 DNA 微阵列杂交进行直接测试。通常使用两种类型的测试：

- **自我测试**。在这个测试中，相同的 RNA 制备物在分别使用 Cy3 和 Cy5 标记的 dUTP 的两个独立反应中标记。然后用混合的标记样品与固定在微阵列上的 DNA 探针杂交。因为实验实际上是样品与自身的对比，阵列上所有单元发出的两种荧光信号的比值应该是恒定的。图 3-14 的散点图是对从自我测试中得到的两种荧光信号的平均强度作图的一个例子。从绝大多数单元发出的两种荧光信号的比值接近于 1。
- **相异的样品对**。在测试标记的可重复性时，通过与 DNA 微阵列的杂交，分析具有不同基因表达谱的成对样品。许多阵列单元显示的两种荧光信号的差别应该很大。当实验重复多次的时候，观察到的两种荧光信号比值的方差可以计算出来。图 3-15 显示了把从两种不同的 RNA 样品制备的 Cy3 和 Cy5 标记的 cDNA 混合在一起，与相同的 DNA 微阵列杂交，对发出荧光信号的比值进行比较的三次实验的散点图。

下列方案是为使用 Cy3 和 Cy5 标记的 dUTP，对按照方案 1 制备的总 RNA 进行标记而设计的。然而，如果使用 1～3 μg poly(A)$^+$或者多核糖体 RNA（方案 5 或方案 6）作为起始材料，则很容易修改以适合于标记纯化的 mRNA。

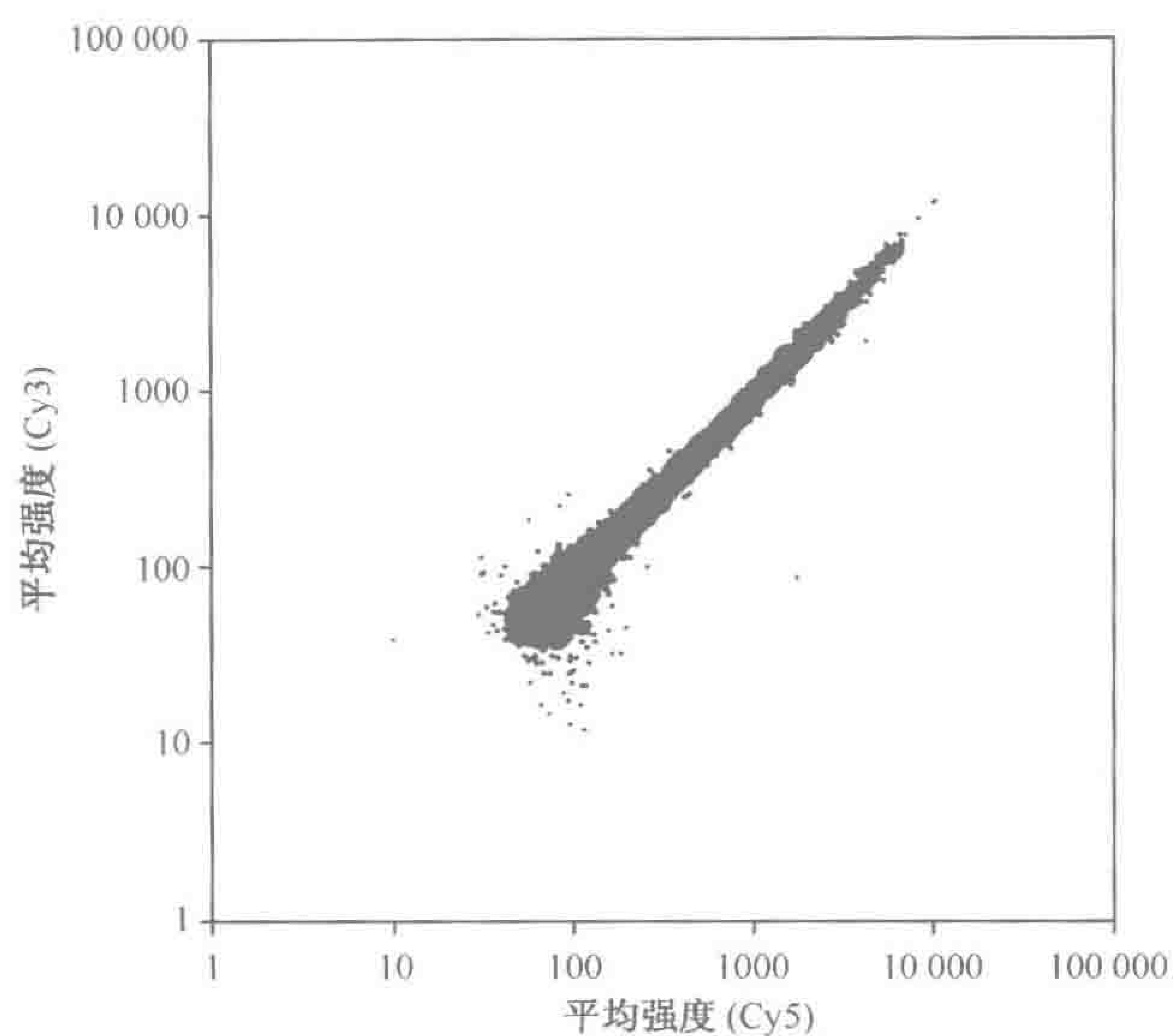


图 3-14 自相关作图

用两种不同的染料标记同一种 RNA 并与具有 6336 种 cDNA 的阵列杂交后，平均荧光强度的对数作图

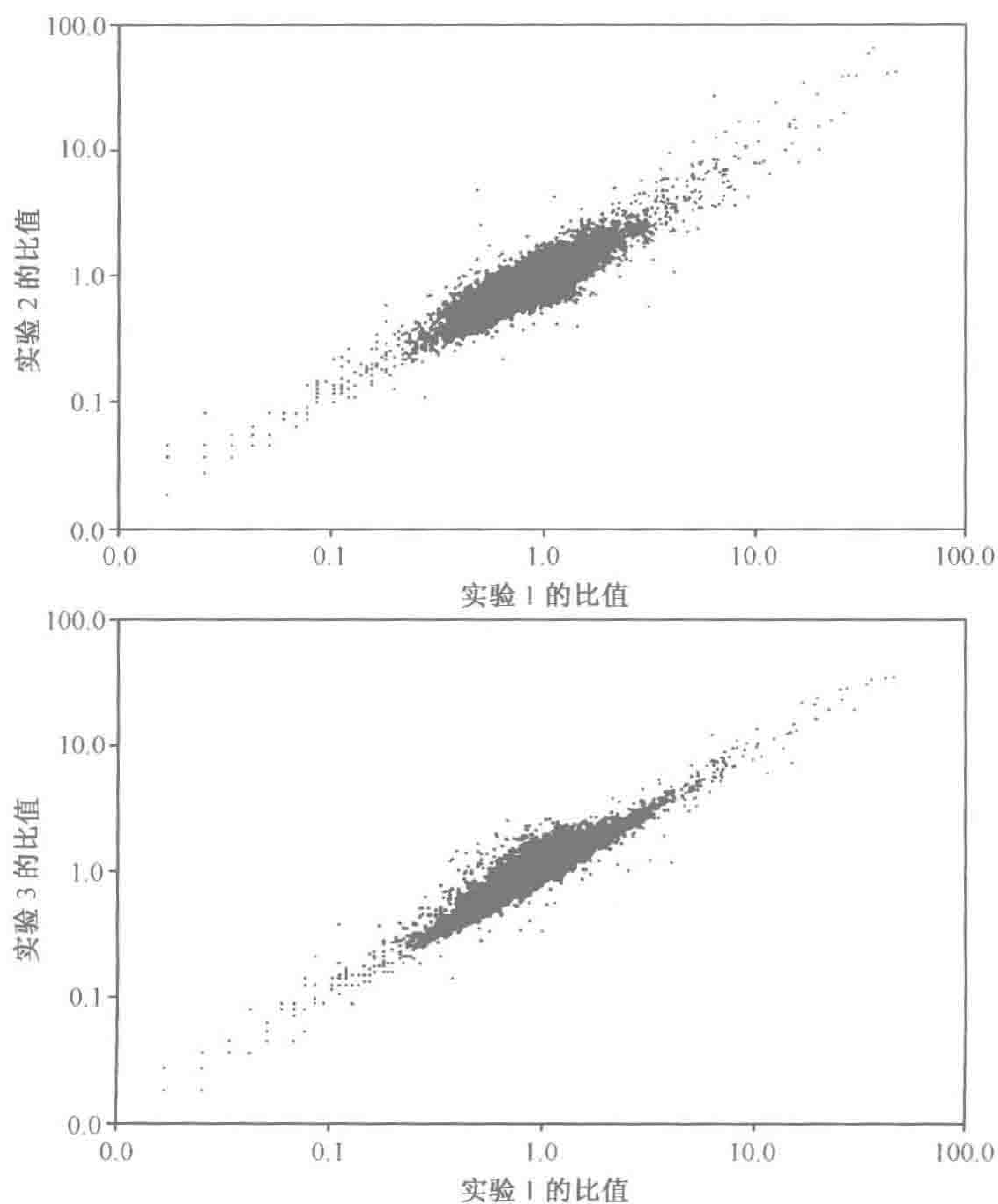


图 3-15 重复性作图

在三次单独的实验中，分别测量两种独立标记的 RNA 与具有 6336 种 cDNA 的阵列杂交得到的平均荧光强度比值。用对数作图显示对通过最小质量限度（minimal quality detection threshold）的测量（3683 个基因）得到的比值进行成对比较的结果

不管使用总 RNA 还是 mRNA，起始材料中应该包含合适的对照 RNA（见第 1 章**点样微阵列中使用的对照**信息栏）。本方案包含四个主要过程：

- 引物与 RNA 的复性。
- 在反转录酶催化下合成 cDNA 第一链。
- 用碱水解 RNA 模板，然后浓缩标记的 cDNA 并除去低分子质量的污染物。
- 检查标记的 cDNA 的质量。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

重要事项：所有缓冲液和溶液使用经 DEPC 处理的 H_2O 配制并分装成小份冰冻保存。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

二硫苏糖醇（DTT）（0.1 mol/L）〈!〉

EDTA(0.5 mol/L，pH 8.0)

NaOH(1 mol/L)〈!〉

每个月新鲜配制并保存在 well-filled 塑料容器中。

50×TAE(Tris-乙酸盐电泳）缓冲液

TE(pH 7.4 和 pH 7.5)

Tris-HCl(1 mol/L，pH 7.5)

酶和缓冲液

cDNA 第一链缓冲液（使用由酶的生产商提供的）

如果没有生产商提供的缓冲液，使用下面的 5×第一链缓冲液：

250 mmol/L Tris-Cl(pH 8.3)

375 mmol/L KCl

15 mmol/L $MgCl_2$

RNasin RNase 抑制剂（30 单位/μl），保存在－20℃。

SUPERSCRIPT Ⅱ RNase H-反转录酶试剂盒（GIBCO)，或者与之相当的

核酸和寡核苷酸

Cy3 标记的 dUTP(1 mmol/L) 和 Cy5 标记的 dUTP(1 mmol/L)(Amersham Biosciences)

结合了羰花菁染料的 dNTP 对光敏感，应该于－20℃下保存在避光的容器中。

如果需要，可以用 Cy3 和 Cy5 标记的 dCTP 代替耦联了荧光生色团的 dUTP。在这种情况下，下述的低浓度 dTTP 混合物应该用低浓度 dCTP 混合物代替。

10×dNTP/低浓度-dTTP 溶液

配制含有 dATP、dCTP、dGTP 的溶液，浓度均为 5 mmol/L，且 dTTP 为 2 mmol/L。将溶液分装成小份于－20℃保存。关于 dNTP 的配置和保存，见附录 2。

Human C_0t-1 DNA（可以从许多供应商得到）

Oligo(dT)$_{12\text{-}18}$

将寡核苷酸以 1 mg/ml 的浓度重悬于 H_2O 中。将溶液分装成小份于－20℃保存。

或者锚定引物（5′-TTT TTT TTT TTT TTT TTT TTT TTV-3′）

将寡核苷酸以 1 mg/ml 的浓度重悬于 H_2O 中。将溶液分装成小份于－20℃保存。

锚定引物是四种合成的寡核苷酸的混合物，由 oligo(dT) 骨架和 3′端的 A、G 或者 C 残基组成。使用锚定引物引导可以确保 cDNA 的合成从 mRNA 的 poly(A) 5′端位置开始。如果微阵列上的杂交探针是对应于转录物的 3′端，建议使用这种引导方法。锚定引物可以从 Sigma 或其他供应商处订购。

总 RNA(＞7 mg/ml)

按照方案 1 所述方法制备。标记效率取决于 RNA 的质量。RNA 的纯度越高，完整性越好，标记的效率就越高。标准的标记反应中包含 40～60 μg 的总 RNA。如果从细胞的收集到 RNA 的纯化所有步骤都进行得很好，少至 25 μg 的总 RNA 就可以用作标记的模板。当 RNA 的供给量不受限制时，每次反应最多可以使用 200 μg 的总 RNA。作为选择，用方案 5 所述方法分离的 mRNA，或者用方案 6 所述方法分离的多核糖体 RNA，也可以用作起始材料。

凝胶

6×贫乏的凝胶上样缓冲液，包含 0.08％(*m/V*) 溴酚蓝〈!〉，0.08％(*m/V*) 的二甲苯青 FF(xylene cyanol FF)〈!〉和 40％的蔗糖，溶解于 H_2O 中。

常用的 2％琼脂糖凝胶（6.0 cm×8.5 cm，带有 2 mm 宽槽），用 1× TAE 缓冲液灌制。

不要向凝胶和电泳缓冲液（running buffer）中加入溴化乙啶。

离心机和转子

台式离心机（用于 1.5 ml 微量离心管的高速 14 000 *g*）

专用设备

荧光扫描仪（如 Molecular Dynamics Typhoon 或者与之相当的设备）

Microcon YM-100(42412 Millipore) 或者 Microcon YM-30(42411 Millipore)

见本方案的步骤 8～10。长度超过 300 个核苷酸的单链 cDNA 大于 90％会被 YM-100 离心装置所保留。对于 YM-30 装置，单链 DNA 名义的截止值约为 50 个核苷酸（见下表）。

性　质	Microcon YM-100	Microcon YM-30
装置的颜色	蓝色	无色
滤过面积/cm^2	0.32	0.32
名义截止值/Da	100 000	30 000
允许的最高离心力/*g*	500,在固定角度转头	14 000,在微量离心机中
离心的时间/min	25℃,5～15 min	25℃,5～8 min
装置的容量/ml	0.5	0.5
保留体积/μl	10	10

PCR 管（0.2 ml，薄壁的）

分光光度计

热循环仪

热循环仪要能容纳 0.2 ml 薄壁 PCR 管，设定为 65℃。

水浴，设定为 42℃和 65℃

方法

引物和 mRNA 复性

1. 准备复性。

如果使用锚定的 oligo(dT) 引物：在 0.2 ml 的薄壁 PCR 管中混合

	为 Cy5 标记加入	为 Cy3 标记加入
总 RNA(>7 mg/ml)	40～60 μg	40～60 μg
锚定引物(2 μg/μl)	1 μl	1 μl
H_2O	至 17 μl	至 17 μl

如果使用 oligo(dT)$_{12\text{-}18}$引物：在 0.2 ml 的薄壁 PCR 管中混合

	为 Cy5 标记加入	为 Cy3 标记加入
总 RNA(>7 mg/ml)	40～60 μg	40～60 μg
oligo(dT)$_{12\text{-}18}$引物(1 μg/μl)	1 μl	1 μl
H_2O	至 17 μl	至 17 μl

2. 将反应混合物在热循环仪中于 65℃下加热 10 min，再于冰上冷却 2 min。

合成第一链 cDNA

3. 向每个复性混合物中加入：

5×第一链缓冲液（SUPERSCRIPT 反转录酶试剂盒）	8 μl
10×dNTP/低浓度-dTTP 溶液	4 μl
Cy5 或 Cy3 标记的 dUTP(1 mmol/L)	4 μl
0.1 mol/L DTT	4 μl
RNasin(30 单位/μl)	1 μl
SUPERSCRIPT Ⅱ(200 单位/μl)	2 μl

反转录酶对在气/液界面的变性非常敏感，因此当混合反应的组分时，要非常小心避免产生泡沫。有时这很难做到，因为许多反转录酶制备物的缓冲液中含有非离子型去污剂，如 Nonidet P-40 和 Triton X-100。

Mo-MLV(SUPERSCRIPT Ⅱ) 对温度敏感，应该储存在－20℃下直到使用前的最后时刻，并且在第一时间放回冰箱。

4. 将反应在 42℃下温育 30 min，然后加入 2 μl 额外的 SUPERSCRIPT Ⅱ。确保

酶和反应物混匀，并在 42℃下再温育 30～60 min。

尽管标记反应最多可以温育 4 h 而不会出现问题，标记的 cDNA 的形成通常在 60～90 min 后就会停止。如果标记 DNA 的产量不是令人满意的，应该根据**测量染料标记 DNA 的效率**信息栏中介绍的方法检查 Cy3 和 Cy5 标记 dUTP 掺入过程随时间的变化。

184 **RNA 的水解**

5. 向每个反应加入 5 μl 0.5 mol/L EDTA。

   加入 EDTA 之后，在继续加入 NaOH（步骤 6）之前，确保反应混合物混匀。核酸会在碱性镁溶液中发生沉淀。

6. 向每个反应中加入 10 μl 1 mol/L NaOH，并将溶液在 65℃下温育 30 min 以水解 RNA。冷却至室温。

   因为 Cy5 遇到碱是不稳定的，用可能的最短时间进行 RNA 的水解。

   两个因素影响反应的成功：

   - **氢氧化钠溶液的纯度**。受到金属污染，或者装在玻璃容器中且暴露于光下的氢氧化物若保存时间过长会产生使 Cy5 降解的溶液，溶液变成黄色。
   - **RNA 的纯度**。受到污染的 RNA 制备物有时包含促进荧光染料降解的物质。

   如果荧光生色团的降解出现问题，可更换 NaOH 溶液，重新纯化 RNA，并尝试将水解时间减少到 20min。

7. 加入 25 μl 1 mol/L Tris-Cl(pH 7.5）中和每个反应。

荧光标记 cDNA 的纯化和浓缩

Microcon YM-l00 超滤管离心超滤被用于纯化和浓缩长度大于 250～300 个核苷酸的荧光标记 cDNA。Cy3 和 Cy5 标记的靶分子可以分开单独处理，或者如果标记的cDNA与相同的微阵列杂交，Cy3 和 Cy5 标记反应（步骤 7）可以混合在一起并同时纯化。当然，如果要从反应混合物中回收未掺入的青色素核苷酸并提取出来（见方案19)，就不应该将反应混合。

较小孔径的超滤管（如 Microcon YM-30）可以用于收集所有长度大于 50～100 bp 的标记产物。见**替代方案：荧光标记的 cDNA 的纯化和浓缩（Microcon YM-30）**信息栏。

8. 转移到 Microcon YM-100 超滤管中：

   中和溶液（步骤 7）

   350 μl TE(pH 7.5)

   20 μg 人的 C_0t-1 DNA

   轻轻吸打将各组分混合，然后将超滤管在 500 *g* 离心 10 min，弃去滤出液。

   这一步骤的目标是将 DNA 脱盐，并除去碱降解的产物（步骤 6)。

   在处理珍贵样品时，向超滤管中加入 200 μl H_2O 并用 500 *g* 离心数秒钟来检查超滤管是否泄漏是值得的。如果过滤器没有正确地密封到超滤管，或者有其他缺陷，水将会流过。如果大部分水被保留下来，可以将水去除并用样品代替。

   在杂交混合物中普遍包含封闭剂，如人的 C_0t-1 DNA 是强制性的。然而，不同的研究人员在不同的阶段加入封闭剂：一些研究人员在合并杂交混合物的时候，而其他人，正如本方案中所描述的那样，更愿意在纯化荧光标记的靶分子时，于一个方便的阶段加入封闭剂。

大量的 DNA 用作载体，减少由于标记的 DNA 结合到表面带来损失。离心前加入封闭剂有助于除去可能污染 C_0t-1 DNA 的盐分和其他杂质。

185

9. 加入 200 μl TE(pH 7.5) 清洗荧光标记的 cDNA 第一链，并将溶液浓缩至体积为约 20～30 μl，通常通过在 500 *g* 离心 8～10 min 实现。

   在通过离心过滤浓缩溶液时，达到希望的最终体积所需要的离心时间是变化的。过长时间的离心会将溶液中的溶剂几乎全部去除。荧光标记的核酸在过滤器上浓缩成这种方式的时候，回收非常困难。为了减少这一问题，用保守估计的离心时间逐渐接近希望的体积。如果有必要，最后的浓缩步骤可以在 SpeedVac 中蒸发液体来实现。真空蒸发，如果没有进行到干燥，不会影响标记 cDNA 的性能。

   在第二次离心之后，保留在过滤器中的标记 cDNA 的颜色通常显著深于流出物，因为它含有更高浓度的吸收光的荧光生色团。

10. 在一个洁净的收集管中倒转 Microcon 超滤管，并在 500 *g* 离心 3 min 以回收浓缩物质。

    在一些情况下，Cy5 标记的 cDNA 被截留在从浓缩溶液中回收的凝胶状的蓝色沉淀中。沉淀的量越大，可以回收的 cDNA 的量就越少。加热可以溶解蓝色凝胶，但这并不能带来好转，因为之后荧光标记的 cDNA 与固定的 DNA 探针均一和非特异性的结合。沉淀的性质是未知的，但它可能包含了很高比例的糖类。Cy5 与浓缩的糖类如糖原具有很高的亲和性。可能的糖类来源有模板 RNA 的制备物或者用作封闭剂的 C_0t-1 DNA。

检查标记 cDNA 的质量

11. 取出 2～3 μl Cy5 标记的 cDNA 用作分析，留下 18～28 μl 用于杂交。将标记的 cDNA 在 4℃下储存于铝箔包裹的试管中。

12. 通过 2％的琼脂糖凝胶电泳分析样品，使用 1×TAE 灌制凝胶和电泳。

    为了获得最高的灵敏度，使用贫乏的凝胶缓冲液（见材料清单），其中包含最少量（平常浓度的 30％）的溴酚蓝和二甲苯青。不要向凝胶和电泳缓冲液中加入溴化乙啶。

13. 用荧光扫描仪（如 Moledular Dynamics Typhoon）扫描凝胶。

    成功的标记产生浓密和扩散的 cDNA，长度为 400～1000 个以上核苷酸，含有最少量的低分子质量转录物（如图 3-13，泳道 C）。标记很弱和（或）显著水平的低分子质量物质是无效标记的指示（如图 3-13，泳道 A）。

    如果需要，标记效率可以使用分光光度法测量。关于详细情况，见**测量染料标记 DNA 的效率**信息栏。

186

替代方案：荧光标记的 cDNA 的纯化和浓缩（Microcon YM-30）

为了纯化和浓缩大小在 50～100 bp 的荧光标记 cDNA 产物，一些研究人员更倾向于使用小孔径的超滤管（如 Microcon YM-30）。

附加材料

Cy3 和 Cy5 标记反应（主要方案的步骤 8）

人的 C_0t-1 DNA(20 mg/ml)（GIBCO）

Microcon YM-30 超离心过滤装置（Millipore）
Poly(A)$^+$RNA(10 mg/ml)(P 9403 Sigma)
TE(pH 8.0)
酵母 tRNA(10 mg/ml)(15401-011，GIBCO)

方法

1. 将 380 μl TE(pH 8.0) 加入到 Microcon YM-30 超滤装置的蓄液池中。
2. 向相同的过滤装置中加入约 60 μl 的 Cy5 标记反应物（步骤 8）和 60 μl Cy3 标记反应物。
3. 将上过样的装置放入生产商提供的一支小管中。
4. 室温下，将装置在微量离心机中以 14 000 g 离心 7～8 min。
5. 从小管中取出流出液体，并将 450 μl 加入装置的蓄液池中。
6. 室温下，将装置在微量离心机中以 14 000 g 离心 7～8 min。
7. 除去流出液体，向蓄液池内管中加入：

TE（pH 8.0）	450 μl
人的 C_0t-1 DNA	20 μg
poly(A)$^+$RNA	20 μg
酵母 tRNA	20 μg

8. 将初始的收集管弃去，以一支新的收集管代替，将超滤装置插入新的收集管中。室温下，将装置在微量离心机中以 14 000 g 离心 7～8 min。

   如果需要，继续离心直到保留体积减少到 5～10 μl。在滤膜的中心部分干燥，靶分子在滤膜的边缘形成环状液体之后，可以得到这样少的体积。确保不要让滤膜完全干燥。
9. 将装置倒转放入洁净的收集管中，以 500 g 离心 2 min 回收荧光标记的 cDNA 。

   在此阶段靶分子通常带紫色。

187

方案 14

用氨基-烯丙基染料对 DNA 进行间接荧光标记

Joseph DeRisi（生物化学与生物物理学系，加州大学，旧金山，加利福尼亚，94143）

荧光标记的 cDNA 通常是在 RNA 模板的反转录过程中掺入结合了染料的核苷酸来产生的。然而，大的荧光染料分子给 RNA 依赖的 DNA 聚合酶带来了严重的空间（排列）问题，以致荧光标记的 dNTP 掺入 cDNA 的效率远低于未取代的 dNTP。这就产生了非均匀标记的低荧光强度 cDNA，降低了低丰度 mRNA 检测的灵敏度，并且增加了重复实验的差异。此外，Cy3-dUTP 和 Cy5-dUTP 掺入到标记 DNA 中的效率差别也减弱了灵敏度。青色素染料耦联的核苷酸掺入的低效和可变带来的金钱代价也是高昂的，

因为荧光标记的核苷酸类似物的价格非常昂贵。很明显，价格低廉、效率更高的标记方法将更受欢迎。

一个非常有吸引力且可行的方法是，通过化学合成或者酶促催化，将一种带有化学反应性氨基的更合适的核苷酸类似物掺入到 cDNA 中，荧光染料可以在之后连接到这种核苷酸上。这一想法由 Randolph 和 Waggoner(1997) 实现的，他们将氨基-烯丙基 dUTP（图 3-16）掺入到 DNA 中，并使得到的伯氨基与 Cy3 的琥珀酰亚胺酯连接结合（图 3-17）。Randolph 的文章清晰而简明地描述了随着荧光基团的密度和连接染料与核苷酸的连接子（linker）长度的变化，氨基-烯丙基的取代和 Cy3 的结合对熔解温度、DNA 的消光系数和染料的量子产率的影响。当标记的密度约为每 6 个核苷酸 1 个荧光分子时，该比例获得了最强的信号要比使用染料标记的核苷酸进行直接标记高。

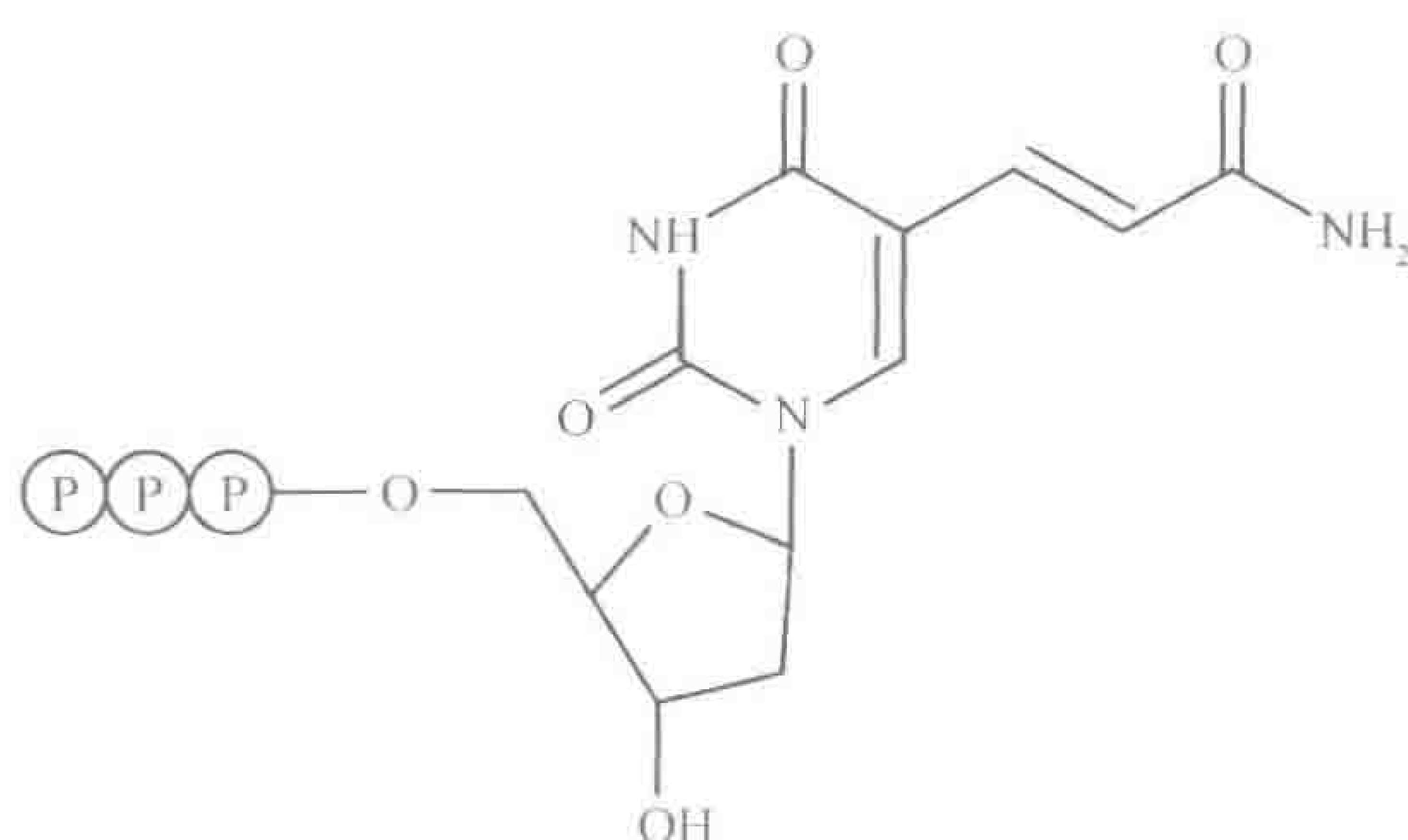


图 3-16　氨基-烯丙基取代的 dUTP

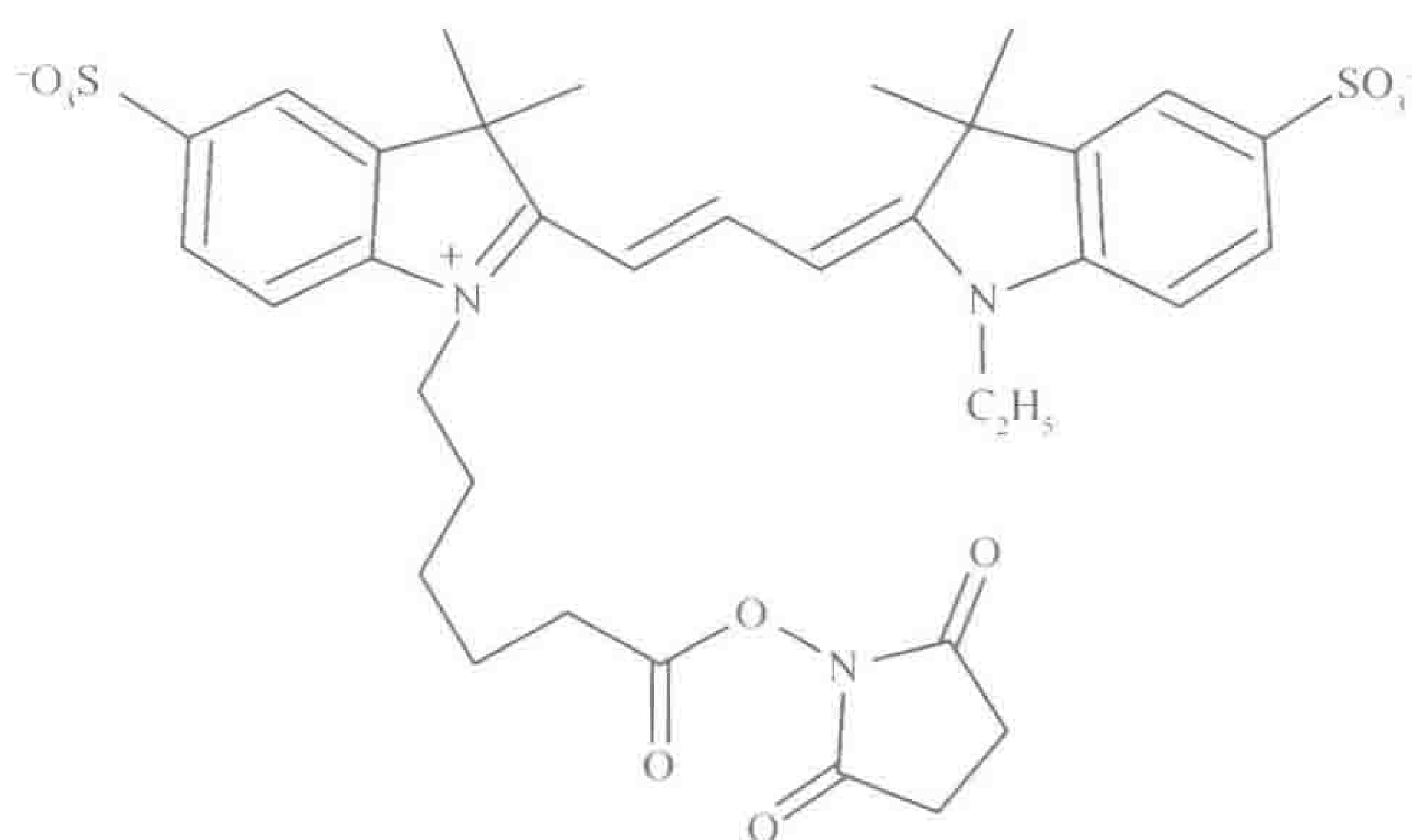


图 3-17　Cy3 的结构

使用氨基-烯丙基染料以更高的染料：核苷酸比例进行标记是可能的，但是会导致荧光信号强度的下降，量子产率也是如此。这种效应是由于非产出性的荧光基团：荧光基团和荧光基团：核苷酸相互作用导致了荧光信号的猝灭。染料和核酸的化学耦联在免疫化学、DNA 测序和原位杂交中已经使用了许多年（Brumbaugh et al. 1988），现在这一技术重新成为微阵列分析中一种广泛使用和充满活力的方法（Hughes et al. 2001）。

与直接的、酶催化的标记方法相比，将反应性染料与氨基取代的核苷酸连接需要花费更长的时间，也需要更多的步骤。然而，它所带来的荧光亮度的优势及没有酶促反应的偏向性和较低的价格，使得这些额外的努力是值得的。另外，现在许多染料都可以得

到 *N*-羟基琥珀酰亚胺酯形式（包括 Cy3 和 Cy5、Alexa 染料、荧光素、bodipy、得克萨斯红、FITC 和罗丹明）。这些酯化的染料（可以从 Molecular Probes 或其他供应商得到）可以和任何带有反应性氨基的核酸分子结合（Panchuk-Voloshina et al. 1999）。

有多种 RNA 依赖的或者 DNA 依赖的 DNA 聚合酶可以催化 5-(3-氨基-烯丙基)-2′-脱氧尿嘧啶 5′-三磷酸盐（aa-dUTP）掺入到 DNA 中。掺入效率与未修饰的 dUTP 几乎完全一致。标记的密度取决于反应混合物中 aa-dUTP∶dTTP 的比例。在 cDNA 合成以后，用碱水解 RNA，然后用游离酸或者缓冲液，如不含伯氨的 HEPES 中和。下面的方案主要根据 Randolph 和 Waggoner(1997) 及 Hughes 等（2001），加入了 DeRisi 和 Bowtell 实验室所作的修改和注释。

189

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

重要事项：所有缓冲液和溶液使用经 DEPC 处理的 H_2O 配制并分装成小份冰冻保存。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

无水二甲基甲酰胺（DMF)〈!〉或二甲亚砜（DMSO)〈!〉

EDTA(1.0 mol/L，pH 8.0)

HEPES(1.0 mol/L，pH 7.4)

NaOH(1.0 mol/L)〈!〉

醋酸钠（3 mol/L，pH 5.2)

碳酸氢钠缓冲液（1 mol/L，pH 9.0)

酶和缓冲液

cDNA 第一链缓冲液（使用由酶的生产商提供的）

如果没有生产商提供的缓冲液，使用下面的 5×第一链缓冲液：

250 mmol/L Tris-Cl(pH 8.3)

375 mmol/L KCl

15 mmol/L $MgCl_2$

RNasin RNase 抑制剂（30 单位/ μl)

保存在－20℃。

SUPERSCRIPT Ⅱ RNase H-反转录酶试剂盒（GIBCO)，或者同类试剂盒（如来自 Stratagene，Inc. 的 StrataScript 反转录酶，或者 CLONTECH 的 PowerScript 反转录酶）

核酸

关于 dNTP 的配置和保存，见附录 2。

dATP、dCTP、dGTP、dTTP（每种的浓度为 100 mmol/L）

dNTP/低浓度-dTTP 溶液

配制含有 dATP、dCTP、dGTP 的溶液，浓度均为 5 mmol/L，且 dTTP 浓度为 2 mmol/L。将溶液分装成小份于－20℃保存。关于 dNTP 的配置和保存，见附录 2。

Human C_0t-1 DNA（可以从许多供应商得到）

Oligo(dT)$_{12\text{-}18}$

将寡核苷酸以 1 mg/ml 的浓度重悬于 H_2O 中。将溶液分装成小份于－20℃保存。

或者锚定引物（5′-TTT TTT TTT TTT TTT TTT TTT TTV-3′）

将寡核苷酸以 1 mg/ml 的浓度重悬于 H_2O 中。将溶液分装成小份于－20℃保存。

锚定引物是四种合成的寡核苷酸的混合物，由 oligo(dT) 骨架和 3′端的 A、G 或者 C 残基组成。使用锚定引物引导可以确保 cDNA 的合成从 mRNA 5′端的 poly(A) 位置开始。如果微阵列上的杂交探针是对应于转录物的 3′端，建议使用这种引导方法。锚定引物可以从 Sigma 或其他供应商处订购。

Poly(A)$^+$、多核糖体 RNA 或扩增的 aRNA（约 200 μg/ml）

按照方案 5 和 7 所述方法制备。

190

随机引物（4 μg/μl，溶解于 TE 中，pH 7.6）

由于具有一致的长度，且没有序列偏向性，随机序列的合成寡核苷酸是选中的引物。最佳长度的寡核苷酸（六聚体和七聚体；Suganuma and Gupta 1995）可以通过商业途径购买（如 Pharmacia 或 Boehringer Mannheim），或者就地在自动合成仪上合成。将引物溶液分装成小份于－70℃下保存。

专用试剂

aa-dUTP[5-(3-氨基-烯丙基)-2′-脱氧尿苷-5′三磷酸钠盐]（Sigma）

将 aa-dUTP 溶解于 H_2O 中，浓度为 100 mmol/L。

Fluorolink Cy3 和 Cy5 Monoreactive Dye——5 个包装（Amersham Biosciences）

关于配制和保存酯化 Cy 染料溶液的说明，见本方案的步骤 13-15。

Amersham 以干燥团块的形式发送青色素染料的琥珀酰亚胺酯，密封在放有少量干燥剂的铝箔袋中。如果干燥剂由深蓝色变成浅粉红色，则潮气可能污染了样品。污染的样品应该退还 Amersham 并得到退款。当羰花青染料的酯暴露到水中时，染料分子上的反应部分迅速水解。必须将冻干的染料于黑暗中保存在干燥器内。

专用设备

台式离心机（高速 16 000 g），适用于 1.5 ml 微量离心管

Microcon YM-30(Millipore)

QIAquick PCR 纯化试剂盒（50）（28104，QIAGEN）

水浴（恒温），加热装置或者热循环仪，设定为 70℃、65℃和 42℃。

方法

通过反转录将 aa-dUTP 掺入到 cDNA 第一链中

1. 将下面的 dNTP 合并，配制 50×dNTP＋aa-dUTP 混合物：

100 mmol/L dATP　　10 μl

100 mmol/L dCTP　　10 μl

100 mmol/L dGTP　　10 μl

100 mmol/L dTTP　　6 μl

100 mmol/L aa-dUTP　　4 μl

将 50×混合物分装成方便使用的量并冰冻保存于－20℃或者－80℃。

标记的强度取决于反应混合物中 dTTP∶aa-dUTP 的摩尔比。在 1∶1（Hughes et al. 2001）和 3∶2（这里使用的）之间的比例通常适合于标记来自酵母或者哺乳动物细胞的 cDNA 第一链。然而，如果标记的强度不够，可以在预实验中使用一系列的比例来确定最佳的 dTTP∶aa-dUTP 的比例。标记密度可以使用**测量染料标记 DNA 的效率**信息栏中介绍的方法来测量。

191

2. 为将引物和 mRNA 复性，混合下面的试剂：

Poly(A)$^+$ RNA 或者多核糖体 RNA（约 200 μg/ml）　　2 μg

引物（1～2 μg/ml）　　1 μl

H_2O　　补充体积至 15.5 μl

当使用纯化的 RNA 作为模板时，cDNA 第一链的合成通常使用 oligo(dT)$_{12-18}$或者一套锚定引物作为引物。然而，随机引物也可以与 oligo(dT)$_{12-18}$或者一套锚定引物共同使用。

3. 将反应混合物在 70℃下孵育 8 min，然后立刻将反应混合物转移到冰浴中。

4. 合并 cDNA 合成的主要混合物：

5× cDNA 第一链缓冲液　　6 μl

50× aa-dNTP　　0.6 μl

DTT(0.1 mol/L)　　3 μl

SUPERSCRIPT Ⅱ RT(500 单位/μl)　　2 μl

RNasin(30 单位/μl)（可选的）　　2 μl

H_2O　　补充体积到 14.5 μl

所示体积是每个标记需要的。

反转录酶对在气/液界面的变性非常敏感，因此当混合反应的组分时，要非常小心避免产生泡沫。有时这很难做到，因为许多反转录酶制备物的缓冲液中含有非离子型去污剂，如 Nonidet P-40 或者 Triton X-100。

Mo-MLV(SUPERSCRIPT Ⅱ）对温度敏感，应该储存在－20℃下直到使用前的最后时刻，并且在第一时间放回冰箱。

5. 向每个延伸反应（步骤 2）中加入 14.5 μl 主要混合物。

6. 将完整的反应混合物在 42℃下温育 2h。

水解和纯化第一链产物

7. 向每个反应混合物中加入 3 μl 0.1mol/L EDTA。

加入 EDTA 之后，在继续加入 NaOH（步骤 6）之前，确保反应混合物混匀，因为核酸会在碱性镁溶液中发生沉淀。

8. 向每个反应混合物中加入 3.5 μl 0.1 mol/L NaOH，65℃下孵育 10 min。

9. 加入 HEPES(pH 7.0) 至终浓度为 500 mmol/L 以中和水解反应。

Tris 带有游离的氨基，不能用来中和，因为它会干扰后续耦联反应。

作为步骤 10～12 中纯化方法的替代，第一链反应产物可以使用 QIAquick 纯化柱（见步骤 12 下面**纯化氨基-烯丙基标记的 cDNA 第一链**信息栏）通过色谱法纯化。然而，QIAGEN 提供的标准洗脱缓冲液中含有 Tris 碱，它带有一个会干扰后续的 aa-cDNA 与酯化的染料耦联的氨基。因此，使用 H_2O 而不是 QIAGEN 的 EB 缓冲液，将 aa-cDNA 从 QIAquick 柱洗脱（见步骤 12 下面的信息栏）。

10. 用 H_2O 将反应体积调整到 500 μl。通过 Microcon YM-30 浓缩器离心浓缩 cDNA产物。

尽量使样品的体积降低至小于 10 μl。在常用的离心机中以最大速度（16 000 *g*）离心6～10 min 可以实现。不要进行旋转干燥，因为这会使得 cDNA 难以回收。

11. 用 H_2O 将浓缩溶液的体积调整到 500 μl 并重复步骤 10 至少两次。

本步骤的目的是除去水解的 RNA、NaOH 和缓冲液成分。

12. 氨基-烯丙基取代的 cDNA(<10 μl) 可以在－20℃长期储存。

纯化氨基-烯丙基标记的 cDNA 第一链

一些研究人员在制备用于标记的 cDNA 第一链时更倾向于将其吸附到层析柱上，然后用乙醇沉淀。QIAquick 柱作为销售的 QIAquick PCR 纯化试剂盒(QIAGEN) 中的一部分，装有依赖于 pH 和盐的结合 DNA 的硅胶膜。按照生产商的说明分别纯化每种标记的样品，进行以下修改。下面的顺序代替步骤 10～12。

a. 向中和的 cDNA 溶液中加入 500 μl(5 倍体积) 的 PB 缓冲液 (QIAGEN)，轻轻涡旋混合溶液。

b. 将一根 QIAquick 柱放入到 2 ml 的收集管中。将 cDNA 溶液加到 QIAquick 柱的中央，并以 10 000 *g* (13 000 r/min) 离心 1 min，弃去流出液体。

c. 向 PE 缓冲液 (QIAGEN) 中加入适量的乙醇 (见瓶上说明)。向 QIAquick 柱中加入 750 μl 含有乙醇的 PE 缓冲液。

d. 10 000 *g* 离心 1 min。弃去流出液体，用 750 μl 含乙醇的 PE 缓冲液再次冲洗。然后再次在 10 000 *g* 下离心 1 min，弃去洗出液体。

e. 将 QIAquick 柱放到 1.5 ml 的微量离心管中。向柱中央加入 30 μl H_2O，静置 2 min，然后在 10 000 *g* 下离心 1 min，收集洗出液体。

重要事项：不能使用 QIAGEN EB 缓冲液洗脱（见步骤 9 注释，上文）。

f. 重复步骤 e，将洗出液体收集在同一支管中。

g. 加入 0.1 体积的 3 mol/L 醋酸钠 (pH 5.2) (约 10 μl)，然后加入 2.5 倍体积乙醇 (约 275 μl)。

有些研究人员加入惰性载体 (如 20 μg 糖原) 以提高沉淀效率。

h. 将溶液在－20℃放置 30 min，然后在 4℃下以 13 000 *g* 离心 10 min 回收cDNA。

i. 用 500 μl (译者注：原书误为 500ml) 冷的 75％乙醇洗涤团块。4℃下以13 000 *g*离心 10 min 回收 cDNA。

j. 仔细吸出乙醇，注意不要影响 cDNA 团块。让 cDNA 在空气中干燥，直到最后一丝可见的乙醇挥发掉。

k. 将 cDNA 溶解在 10 μl H_2O 中并直接进行步骤 13，或者将 DNA 储存在－20℃直到需要用时。

有些研究人员更倾向于在步骤 f 后将样品真空干燥，而不使用乙醇沉淀它们（步骤 g～j）。

Cy 染料酯化物的分装

193

13. 将 Cy-染料酯化物固体团块重悬于 12 μl 无水 DMF 或 DMSO 中。
14. 因为单管中的酯化染料通常足以标记至少 12 个样品，将重悬的染料按照 1 μl 的分量分装到带旋盖的试管中。用 SpeedVac 旋转浓缩器将分装的染料干燥，不要加热。
15. 将干燥的分装染料保存在 4℃下，放入避光的盒中，最好置于真空中并有大量的干燥剂。

    这样储存可以避免染料被潮气污染。

氨基-烯丙基取代的 cDNA 与 N-羟基琥珀酰亚胺酯化的染料偶联

16. 向标记的 cDNA（来自步骤 12，10 μl）中加入 1.5 μl 新鲜配制的 1 mol/L 碳酸氢钠缓冲液（pH 9.0）。
17. 取出一份分装后保存的 N-羟基琥珀酰亚胺酯染料（步骤 15），并进行剧烈上下吹打，用碳酸氢钠缓冲的 cDNA 溶液重悬团块。
18. 室温下，在黑暗环境中将耦联反应温育至少 60 min。
19. （可任选的）有些研究人员在继续 QIAquick 柱纯化（步骤 21）之前加入羟胺至终浓度为 1.5 mol/L 来猝灭耦联反应。然而，其他人（包括 DeRisi 实验室）认为羟胺猝灭是没有必要的，并且可能是有害的。

去除游离的酯化染料

20. 按照步骤 12 下面信息栏中的步骤 a～g，通过 QIAquick 柱色谱分别纯化每个染料标记的样品。
21. 将步骤 g（见信息栏）中得到的最终洗出液体用 Microcon YM-30 浓缩器或者真空干燥器浓缩到希望的体积。
22. 将标记好的 cDNA 在 4℃下储存于铝箔包裹的试管中。使用凝胶电泳分析标记好的 cDNA 的样品并按照方案 13 中的步骤 12 和步骤 13 介绍的方法计算标记效率。

    一个好的探针应当在每 75 个核苷酸中含有 1 个以上的染料基团。

194

方案 15

用于为筛选 DNA 微阵列产生靶序列的 mRNA 制备物的扩增

James Eberwine[*,†]，**Paolo Marciano**[‡]，**Tracy McIntosh**[‡]，和 **Janet Estee Kacharmina**[*,†]（[*]药理学、[†]精神病学、[‡]神经外科学系，宾州大学医学中心，费城，宾夕法尼亚，19104）

用 DNA 微阵列来检测和测量基因表达谱的变化时，标记的靶序列应该忠实地反映原始组织或细胞中 mRNA 的丰度。当荧光标记的靶序列为直接从 mRNA 模板拷贝而来时，产生严重偏差的机会最小，例如，在单个酶促反应中，RNA 依赖的 DNA 聚合酶

将 mRNA 转录为荧光标记的 cDNA 第一链。然而，将 mRNA 直接转变为标记的 cDNA 需要大量的模板，例如，从显微切割组织或是使用荧光激活的细胞筛选方法从少量细胞中得到的稀有细胞，就不能满足这个要求。在这些情况下，制备的 mRNA 必须扩增到一定的数量，足以用来合成足够数量的靶序列。在本方案中，我们讨论常用的反义 RNA(antisense RNA，aRNA）扩增方法来产生用于靶序列合成的模板。我们介绍少量的 RNA（约 1 ng 总 RNA）如何拷贝成 cDNA，并且通过使用噬菌体 T7 DNA 依赖的 RNA 聚合酶进行体外转录扩增。对石蜡包埋的固定组织进行显微切割得到的 RNA 的分离和扩增方法将在方案 11 和第 5 章中介绍。

扩增由 RT-PCR 产生的 cDNA

在该方法中（图 3-18），使用 5′端带有特定任意序列（UAR）的 oligo（dT）分子作为引物，将 mRNA 拷贝为 cDNA 第一链。第一链的合成结束后，在末端脱氧核苷酸转移酶（terminal deoxynucleotide transferase，TdT）催化作用下向 cDNA 第一链的 3′端加上 dG 残基。同聚的 dG 3′端尾部和合成的 oligo(dT）引物 5′端特定的序列接下来被用于由 *Taq* DNA 聚合酶催化的标准 PCR 反应中引物的复性位点。这类 PCR 是相对高效的，它们会改变原始 mRNA 的相对丰度。在 PCR 过程中，因为长度和二级结构的差异，有些cDNA会比其他的更有效地扩增。例如，在测量两种cDNA相对丰度的实

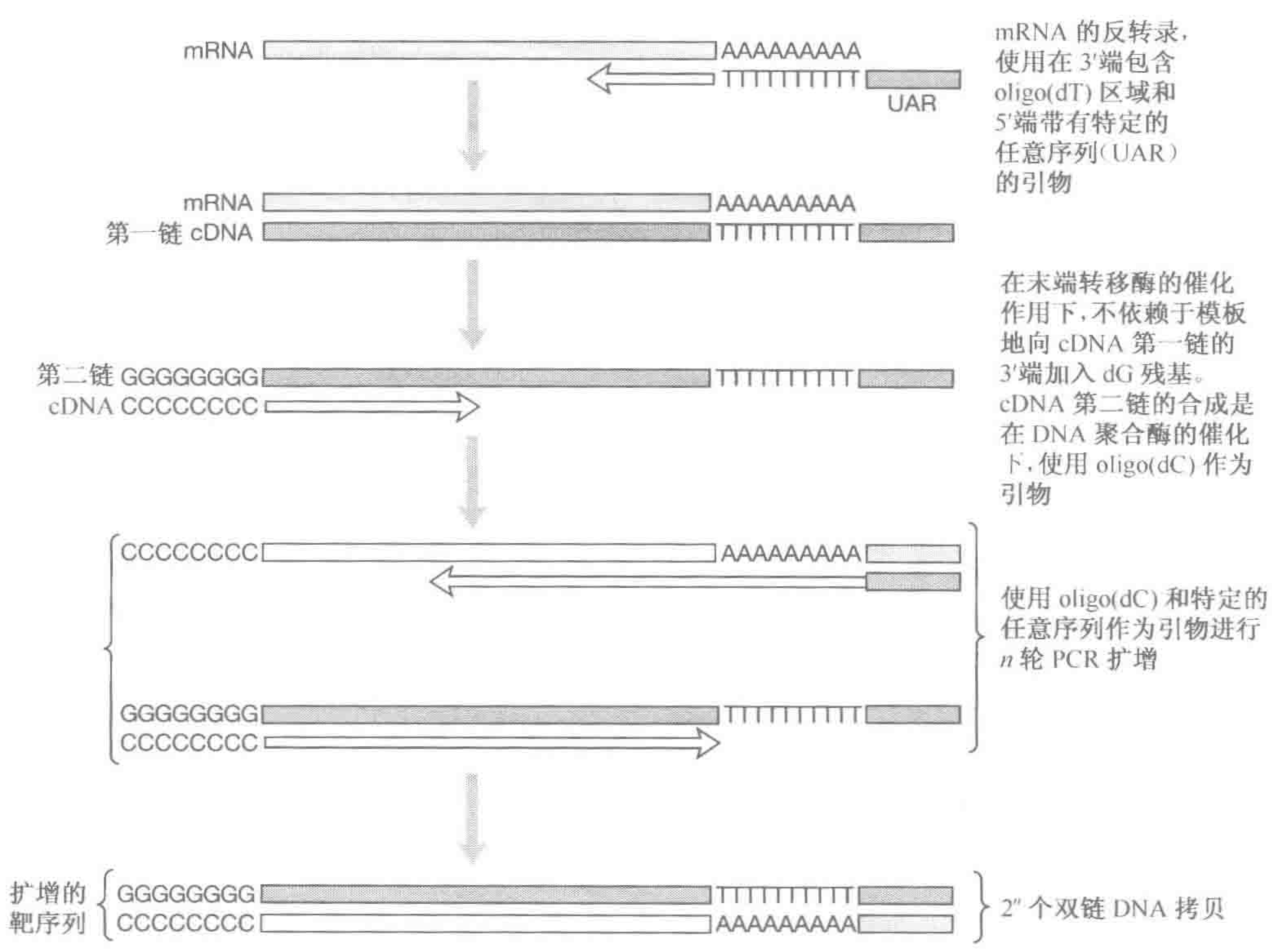


图 3-18 RT-PCR 示意图

显示了对一类 mRNA 进行 PCR 扩增包括的主要步骤。对一种特定mRNA种类的任何正、负偏差在这种方法中都被指数放大，导致靶序列不能准确地反映出初始mRNA中每种 mRNA 的丰度

验中，如果 cDNA 1 的扩增效率是 cDNA 2 的 90%，那么在经过一轮 PCR 后 cDNA 1 比 cDNA 2 的摩尔比例就会产生 9/10 的偏离。

如果 PCR 重复 35 次循环，而扩增效率的差异保持不变，在 35 轮的 PCR 之后，cDNA 1 相对于 cDNA 2 的百分比为最初量的 $(9/10)^{35}$，即 0.028%。因此，在多轮扩增后产生的 cDNA 与产生它们的 mRNA 几乎没有相似之处。由于潜在的强烈偏差，用多轮 PCR 产生的靶序列来进行微阵列筛选是不合适的。

用 T7 RNA 聚合酶扩增 aRNA

由 Van Gelder 等在 1990 年首次提出的 aRNA 方法已经广泛应用于各种目的研究中，包括产生用于微阵列分析的靶分子。aRNA 方法使用噬菌体编码的 RNA 聚合酶将 DNA 模板转录成 RNA（图 3-19），这是通过使用 5′端编码合成的 T3、SP6、T7 DNA 依赖的 RNA 聚合酶的启动子的 oligo(dT) 或者随机六聚体引物来合成 mRNA 的单链 cDNA 拷贝而实现的。

196

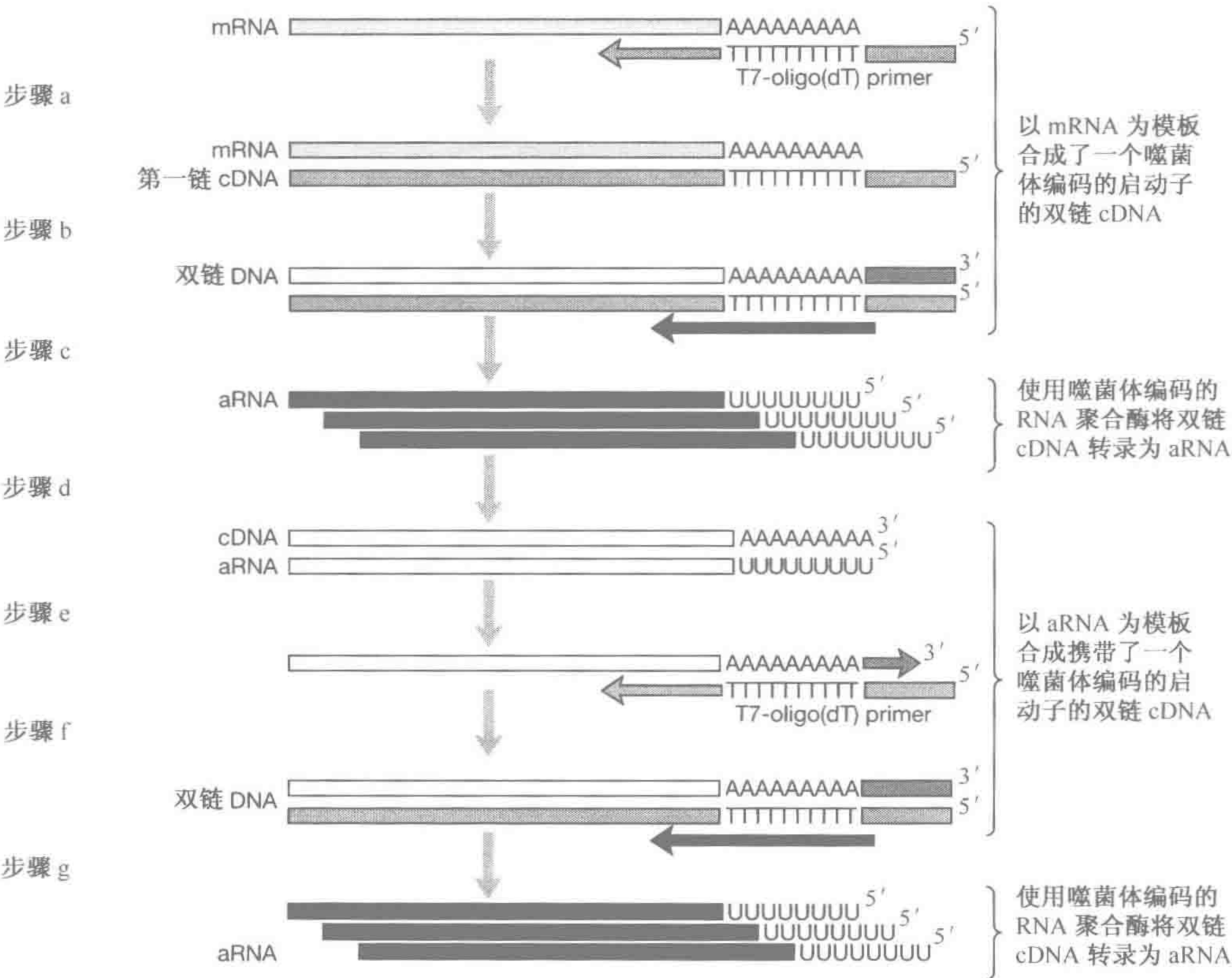


图 3-19　aRNA 的合成

该示意图显示了 mRNA 的反义 RNA 拷贝合成所包括的步骤。图中显示了两轮 aRNA 的合成（步骤 a～c 和步骤 d～g）。第一轮的 aRNA 产物用作第二轮合成的模板。方案中的步骤总结在下表中。关于进一步的详细情况，见方案

步　骤	模　板	引　物	酶	产　物
a	mRNA	T7-oligo(dT)	反转录酶	mRNA-cDNA 杂合子
b	cDNA	随机六聚体	*E. coli* DNA 聚合酶 Ⅰ	双链 cDNA
c	双链 cDNA	无	T7 RNA 聚合酶	aRNA
d	aRNA	随机六聚体	反转录酶	aRNA-cDNA 杂合子
e/f	cDNA	T7-oligo(dT)	E. *coli* DNA 聚合酶 Ⅰ	双链 cDNA
g	双链 cDNA	无	T7 RNA 聚合酶	aRNA

将单链 DNA 转化成双链 cDNA 后产生了可以用适当的 DNA 依赖的 RNA 聚合酶转录成 aRNA 的模板。这种方法的主要特点是：①aRNA 按照线性速度的方式积累（相对于 PCR 中的指数速度）；②aRNA 不再用作反复扩增循环的模板。由于在线性扩增过程中，mRNA 丰度的偏离线性扩增减至最小，最终的 aRNA 靶序列准确的代表了原始的所有 mRNA 中各种 mRNA 分子的大小、复杂度和丰度（Van Gelder et al. 1990；Eberwine et al. 1992）。总 RNA 或者是 mRNA 都可以用作起始模板。

该方法优化以后，体外合成的 aRNA 的数量要比最初 mRNA 的量高出 1000～2000 倍。如此高效的扩增使得几纳克的起始 RNA 就可以在 DNA 微阵列上进行表达分析。如果 mRNA 的量更少，需要再进行一次扩增。连续两次扩增后，理论上可以对起始 mRNA 进行几百万倍的扩增。实际应用中，两轮扩增后 aRNA 的数量增加大约 10^5 倍，并且几乎没有增加随机的或者系统性偏差的污染（Wang et al. 2000）。通过两轮扩增产生足够量的 aRNA 使得分析从单个细胞中得到的微量 mRNA(0.1～1.0 pg）成为可能（Kachar mina et al. 1996；Crino and Eberwine 1997）（见方案 11）。

体外合成时，荧光标记的核糖核苷酸掺入 aRNA 分子的效率很低。三种方法可以解决这个问题并且从 aRNA 产生靶序列。

- **将生物素标记的核糖核苷酸掺入到 aRNA 链中**。生物素化的 aRNA 靶序列与 DNA 微阵列杂交后，使用荧光素标记的亲和素或者抗生物素抗体来检测生物素。关于靶序列制备的实验细节，见方案 18。
- **使用反转录酶和 5′端带有荧光标记的随机引物，将 aRNA 产物转化为 cDNA**。5′端标记的 cDNA 用作和微阵列杂交的靶序列。
- **将荧光标记的脱氧核苷酸掺入反转录酶催化下，以 aRNA 为模板，随机引物引导合成的 cDNA 中**。然后内部标记的 cDNA 用作靶序列。这实际上是靶序列合成的标准方法（方案 13），除了 aRNA 代替 mRNA 或者总 RNA 作为模板。本方案描述了靶序列制备的实验细节。

从少量的总 RNA 或者 poly(A)$^+$RNA 合成 aRNA 和双链 DNA

本方案产生足够量的 aRNA 模板，使得可以从极少量（约 1 ng）的总 RNA 或者 poly(A)$^+$RNA 中产生荧光标记的靶序列。可以获得大量的 RNA[大于 50 μg 总 RNA 或者大于 2 μg poly(A)$^+$RNA] 时，不需要进行模板的扩增，可以使用方案 13 和方案 14 描述的常用方法进行标记靶序列的合成。该方法包含三组酶促反应：

- 将 RNA 转化为双链 cDNA（步骤 1～4)。
- 将双链 cDNA 转化为 aRNA（步骤 5～7)。

• 将 aRNA 转化为双链 cDNA（步骤 7～14）。

双链 cDNA 可以用作第二轮 aRNA 合成的模板，或者可以按照**附加方案：标记 aRNA**的信息栏中描述的方法通过随机引物标记掺入荧光核苷酸。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

重要事项：所有缓冲液和溶液使用经 DEPC 处理的 H_2O 配制并分装成小份冰冻保存。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度，并过滤灭菌。

二硫苏糖醇（DTT）(0.1 mol/L)〈!〉

乙醇（无水）

据报道，有些批次的乙醇（100%或 200 标准强度）中含有会使微阵列试验中产生很深荧光背景的污染物质，95%的乙醇不存在这个问题。为避免这个问题，可以用 95%的乙醇或已鉴定证明不会产生高荧光背景的无水乙醇。Warner-Graham 公司（Cockeysville，Maryland）出售的无水乙醇（USP 级乙醇，64-17-5）效果很好。

线性聚丙烯酰胺

酚：氯仿（75：25）〈!〉

醋酸钠（3 mol/L，pH 7.0）

酶和缓冲液

噬菌体 T4 DNA 连接酶

噬菌体 T4 DNA 聚合酶（2500 单位/ml）

噬菌体 T7 DNA 依赖的 RNA 聚合酶（1000 单位/μl）

E. coli DNA 连接酶（1000～4000 单位/ml）

E. coli DNA 聚合酶 Ⅰ(10 000 单位/ml)

E. coli RNase H(1000 单位/ml)

10×第一链缓冲液

500 mmol/L Tris-Cl(pH 8.3)

1.2 mol/L KCl

100 mmol/L $MgCl_2$，

反转录酶

重组鼠反转录酶（RT）可以从许多制造商得到。我们强烈推荐使用没有 RNase H 活性的突变形式。在本方案中使用的是 SUPERSCRIPT Ⅱ反转录酶（Invitrogen/GIBCO）。然而，本方案很容易修改以适用于其他商品化的 RNase H^-形式的 Mo-MLV RT。反转录酶的特异活性取决于用于在 *E. coli* 中表达酶的特定 cDNA 克隆并且各个制造商之间存在差异。按照手边各个制造商的说明使用酶（如单位/μg 加入的 RNA 和温育温度）。Mo-MLV RT 对温度敏感，必须保存在－20℃直到需要时。

10×RNA 扩增缓冲液

400 mmol/L Tris-Cl（pH 7.5）

70 mmol/L $MgCl_2$

100 mmol/L NaCl

20 mmol/L 亚精胺

加入亚精胺之前需要过滤。

RNA 酶抑制剂

RNase 的蛋白抑制剂可以与大多数 RNase 结合并抑制其活性，但是不影响野生型 Mo-MLV RT 的活性。许多制造商销售这些抑制剂，商品名也不相同（如 Promega 的 Rnasin 和 5-Prime→3-Prime 的 Prime Inhibitor）。

10×第二链缓冲液

1 mol/L Tris-Cl（pH 7.4）

200 mmol/L KCl

100 mmol/L $MgCl_2$，

400 mmol/L$(NH_4)_2SO_4$

核酸和寡核苷酸

含有四种 dNTP 的 dNTP 溶液，每种浓度为 10 mmol/L

将溶液按照小份分装后保存在－20℃。关于配制和保存 dNTP 的说明，见附录 2。

含有四种核糖核苷酸三磷酸盐的 NTP 溶液，每种浓度为 2.5 mmol/L

将溶液按照小份分装后保存在－20℃。关于配制和保存 NTP 的说明，见附录 2。

Oligo$(dT)_{24}$-T7 启动子

5′AAA CGA CGG CCA GTG AAT TGT AAT ACG ACT CAC TAT AGG CGC$(T)_{24}$

该启动子序列比由 Dunn 和 Studier(1983）定义的 T7 启动子的公认序列要长得多，New England Biolabs 和 Stratagene(TAAT ACGACTCACTATAGGGAGA）有售。在这里定义的寡核苷酸，T7 启动子嵌入在序列中，在启动子的 5′侧翼区域，为 T7 RNA 聚合酶结合提供了空间，启动子 3′侧翼区域的三个核苷酸可以增加 T7 的酶活性。将合成的寡核苷酸溶解在水中，浓度为 100 pmol/μl（约 2μg/μl）。

合成的寡核苷酸要使用聚丙烯酰胺凝胶电泳进行纯化（Sambrook and Russell 2001），并以约 25 pmol/μl（约 0.5 μg/μl）的浓度溶解于 H_2O 中。

随机引物

由于具有一致的长度，且没有序列偏向性，所以随机序列的合成寡核苷酸是有效地将 aRNA拷贝为 cDNA（本方案的步骤 8～10）首选的引物。最佳长度的寡核苷酸（六聚体、七聚体或八聚体；Suganuma and Gupta 1995）可以通过商业途径购买（如 Pharmacia 或 Boehringer Mannheim）或者就地在自动 DNA 合成仪上合成。将引物按照浓度为 1.0 μg/ml 溶解在 TE(pH 7.6）中。引物溶液分装成小份在－20℃下保存。

总 RNA(1 μg/ml）或 poly$(A)^+$RNA（1～10 ng）使用方案 1～10 中介绍的方法之一从合适的来源纯化。

专用设备

Microcon YM-50 柱，或者与之相当的

水浴，设定为73℃、70℃、42℃和16℃

方法

cDNA 第一链和第二链的合成

1. 为合成 cDNA 第一链，在微量离心管中混合以下试剂：

1 ng/μl 总 RNA	1 μl
100 pmol/μl oligo(dT)$_{24}$-T7 启动子	1 μl
H_2O	加至 13 μl

200

密封反应管，将混合物在70℃下温育10 min。从70℃水浴中取出试管，在室温下放置5 min，然后加入：

10×第一链缓冲液	2 μl
0.1 mol/L DTT	1 μl
2.5 mmol/L dNTP 溶液	2 μl
RNase inhibitor	1 μl
SUPERSCRIPT Ⅱ(200 单位)	1 μl

轻轻涡旋混合试剂并将反应混合物短时离心除去气泡。将密封的微量离心管在42℃温育1h。

2. 为合成第二链，向第一链反应中加入以下试剂：

10×第二链缓冲液	10 μl
2.5 mmol/L dNTP 溶液	10 μl
E. coli DNA 聚合酶 Ⅰ(40 单位)	4 μl
E. coli RNase H(1 单位)	1 μl
H_2O	加至 80 μl

轻轻涡旋混合试剂并将反应混合物短时离心除去气泡。将密封的微量离心管在16℃温育2 h。

在第二链合成时，*E. coli* DNA 连接酶用于封闭 DNA 上相邻的5′和3′羟基末端。为了达到这个目的，来自细菌的酶要比噬菌体 T4 DNA 连接酶更为合适，因为它不会将在 cDNA 第二链合成过程中产生的相邻的 DNA 和 RNA 片段相连接（Oakayama and Berg 1982）。

3. 加入2 μl(5单位）噬菌体 T4 DNA 聚合酶并将反应混合物在16℃下温育10 min。

T4 DNA 聚合酶用于补平完成的双链 DNA 的末端。

4. 用150 μl 酚∶氯仿（75∶25）抽提反应混合物。将试管短时离心进行分相，并将130 μl 水（上层）相转移到新的微量离心管中，将试管保存在工作台上。为了进一步纯化双链 cDNA 产物：
   a. 向含有酚∶氯仿的试管中加入100 μl H_2O。
   b. 通过涡旋混合两相并将试管短时离心进行分相。
   c. 将120 μl 水相转移到含有来自第一次抽提时的130 μl 水相的微量离心管中。
   d. 向合并的水相中加入250 μl H_2O。
   e. 将合并的水相转移到一个离心柱，如 Microcon YM-50 中。

f. 在室温下，将柱子以 13 000 *g* 离心 5 min。

g. 用 500 μl H_2O 洗涤柱子并按照步骤 f 进行离心。

h. 在一个新的微量离心管中倒转柱子。将试管在 1000 *g* 下离心 3 min 收集样品（约 15 μl）。

将双链 cDNA 转录成 aRNA

201

5. 将14 μl cDNA 制备物（步骤 4）转移到新的离心管中并加入以下试剂：

试剂	体积
10× RNA 扩增缓冲液	2.0 μl
0.1 mol/L DTT	1.0 μl
2.5 mmol/L NTP 溶液	2.0 μl
RNasin(20 单位)	0.5 μl
T7 RNA 聚合酶（1000 单位/μl）	1.0 μl

轻轻涡旋混合试剂并将反应在 37℃温育 4 h。

6. 向反应中加入 20 μl H_2O，随后加入 40 μl 酚：氯仿（75：25）。将试管短时离心进行分相，并将 30 μl 水（上层）相转移到新的微量离心管中，将试管保存在工作台上。为了进一步纯化 aRNA 产物：

a. 向含有酚：氯仿的试管中加入 20 μl H_2O。

b. 通过涡旋混合两相并将试管短时离心进行分相。

c. 将 20 μl 水相转移到含有来自第一次抽提时的 30 μl 水相的微量离心管中。

d. 向合并的水相中加入 5 μl 3 mol/L 醋酸钠（pH 7.0），随后加入 8 μl 0.25% 线性聚丙烯酰胺和 130 μl 乙醇。

e. 将溶液在－20℃下储存 4～8 h。

f. 4℃下以 15 000 *g* 离心 10 min，回收 aRNA。

g. 轻轻吸去乙醇并让团块在室温下干燥 15～20 min。

7. 将 RNA 团块溶解于 12.5 μl H_2O 中。

将 aRNA 转化为双链 cDNA 用于再次扩增

8. 在纯化的 RNA 中，加入 1 μl 浓度为 1 μg/μl 的随机引物并将溶液在 70℃下温育 10 min。将溶液在冰上快速冷却，然后在室温下保存 10 min。

9. 为合成第一链的，加入下列试剂：

试剂	体积
10×第一链缓冲液	2 μl
0.1 mol/L DTT	2 μl
10 mmol/L dNTP 溶液	0.5 μl
RNase 抑制剂	1 μl
SUPERSCRIPT Ⅱ(200 单位/μl)	1 μl

将各组分混匀并将溶液（最终体积约 20 μl）在 37℃温育 1 h。

10. 将反应混合物在 94℃加热 2 min 然后在冰上迅速冷却。

11. 加入 1 μl 的 oligo$(dT)_{24}$-T7 启动子（100 pmol/μl）并将反应混合物在 70℃温育 5 min，在 42℃下温育 3 min。

202

12. 为合成第二链，加入下列组分：

10×第二链缓冲液	15 μl
10 mmol/L dNTP	7.5 μl
E. coli DNA 聚合酶 I（40 单位）	4 μl
E. coli RNase H(1 单位）	1 μl
H_2O	加至 100 μl

将溶液混合并在 16℃下孵育 2 h。

13. 加入 2 μl T4 DNA 聚合酶并在 16℃孵育 10 min。

    T4 DNA 聚合酶用于补平完成的双链 DNA 的末端。

14. 按照步骤 4 所述纯化得到的双链 cDNA。

15. 为生成第二轮 aRNA，重复步骤 5～7 来扩增这里的双链 cDNA。

    假定 mRNA 占总 RNA（译者注：原书误为总 DNA）的大约 1%；因此对 1 ng 总 RNA 10^3 倍的全面扩增应该产生约 10 ng 的双链 DNA（译者注：原书误为双链 RNA）。

    第二轮 aRNA 产物的质量和大小可以使用琼脂糖-甲醛凝胶电泳检测（Sambrook and Russell 2001）。

    cDNA 即可用于使用下面附加方案中所述的方法之一进行标记。

203

附加方案：标记 aRNA

取决于 aRNA 靶序列如何使用，第二轮 aRNA 的 T7 体外转录产物可以使用多种方法进行标记。对于使用^{33}P 标记靶序列的微阵列，可使用第 4 章方案 2 中所述的方法在 aRNA 合成的同时将^{33}P 标记的核糖核苷酸掺入链中。一种成功用于筛选微阵列的荧光检测方法是基于使用来自 Enzo Diagnostics，Inc 的 RNA 标记试剂盒。在 RNA 聚合酶催化下，荧光 NTP 掺入 RNA 产物的效率通常很差，Enzo 试剂盒通过采用生物素化的核糖核苷酸来促进 RNA 合成而克服这一问题。在与阵列杂交后，使用与生物素有很高亲和力的荧光标记的抗体或亲和素来检测生物素标记的 RNA。本方案中介绍了使用 Enzo 试剂盒合成生物素化的 aRNA。作为替代，aRNA 可以反转录成 cDNA，使用随机引物将 Cy3 或 Cy5 标记的核苷酸掺入 cDNA 产物。

附加材料

RNA 标记试剂盒（Enzo Diagnostics，Inc.；www.EnzoBio.com）

试剂盒成分包括生物素标记的核糖核苷酸和 HY 反应缓冲液

RNeasy 柱（QIAGEN）

可选的，见步骤 2

方法

1. 将下列试剂加入含有第一轮或者第二轮 aRNA 扩增后产物的试管中（主要方案步骤 15）。

10×HY 反应缓冲液	4 μl
生物素标记的核糖核苷酸	4 μl

0.1 mol/L 二硫苏糖醇	4 μl
RNase 抑制剂	4 μl
T7 RNA 聚合酶（1000 单位/μl）	2 μl
H_2O	加至 40 μl

混合试剂并将反应混合物在 37℃下温育 4 h。

2. 使用主要方案步骤 6 描述的酚：氯仿抽提或者乙醇沉淀方法纯化标记的 cRNA 产物，或者使用 QIAGEN 的 RNeasy 纯化柱方法（详细情况见制造商的说明或者方案 1）。

方案 16

204

用于表达分析的少量 mRNA 的扩增

Ena Wang，Francesco M. Marincola（HLA 与免疫遗传学实验室，国家癌症研究所，国立卫生院，贝塞斯达，马里兰，20892）

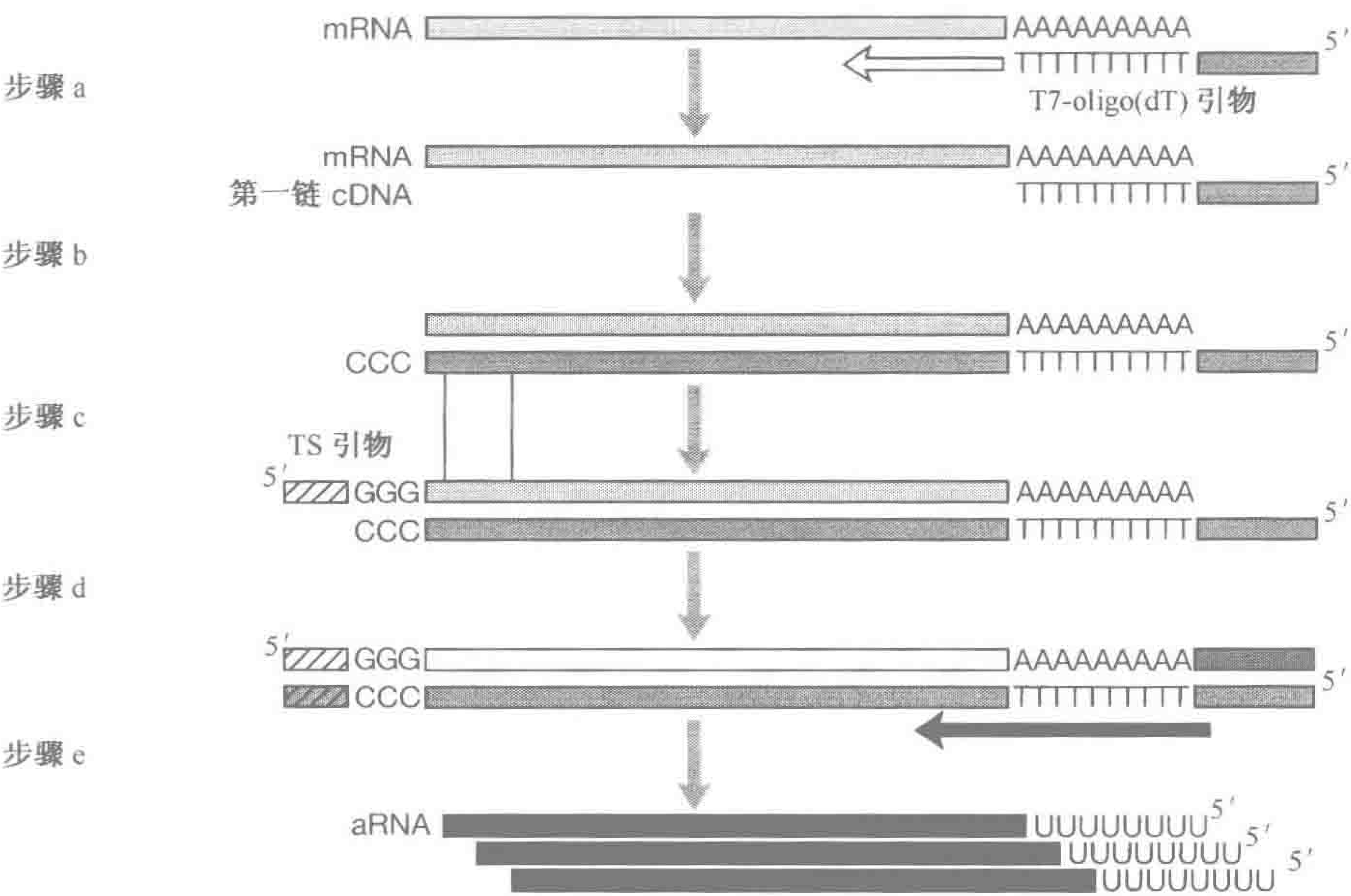
许多生物学过程发生在细胞间隔中，由于空间太小而不能产生足够量的 RNA 用于 DNA 微阵列的表达分析。通过增加荧光强度来试图扩展 cDNA 微阵列的灵敏度（Chen et al. 1998；Zejie et al. 1998；Rajeevan et al. 1999）只取得有限的改进。因此，直到目前为止，微阵列技术还局限于可以抽提大量 RNA ［约 50～200 μg 总 RNA 或者 2～5 μg $poly(A)^+$ RNA］的系统。然而，在过去的数年中，发展了体外扩增 mRNA 或 cDNA 的方法（Lockhart et al. 1996；Phillips and Eberwine 1996；Trenkle et al. 1998），有潜力使得微阵列分析需要的总 RNA 或 $poly(A)^+$RNA 的数量可以减少几个量级。尽管这些扩增方法得到了广泛的接受（Van Gelder et al. 1990；Eberwine et al. 1992；Lockhart et al. 1996；Kacharmina et al. 1999；Luo et al. 1999），但迄今为止几乎很少有系统的工作来评价其局限性，或者是比较已有的实验方案的效率。在 Eberwine 的工作的基础上，我们最近介绍了一种改进的方案（Wang and Marincola 2000），通过将 aRNA 扩增（Phillips and Eberwine 1996）和模板转换（template switching）（Matz et al. 1999）结合，优化了复杂的总 mRNA 的扩增。从极少量的 mRNA（正常加入量的 1/100 000～1/10 000）高保真扩增 aRNA 产生的表达谱与通常使用 $poly(A)^+$ RNA 或者总 RNA 的阵列得到的结果相当（Wang et al. 2000），表明对该方法进行了验证。在本方案中，我们简要地评述 RNA 的扩增方法并在我们的扩增方法中详细描述。

扩增总的 mRNA 而不偏离其相对丰度仍然是一个研究热点。所研究的线性扩增方法在理论上应该保持每种 mRNA 在原始样品中的比例。体外转录带有噬菌体 T7 启动子的双链 cDNA（Van Gelder et al. 1990）提供了一种有效扩增 mRNA 序列的方法（aRNA），并由此产生了用于合成荧光标记的单链 DNA 的模板（Van Gelder et al. 1990；Eberwine et al. 1992；Lockhart et al. 1996；Phillips and Eberwine 1996；Kacharmina et al. 1999；Luo et al. 1999）。然而，对这些方法的验证主要是依赖于使用

northern 分析或者原位杂交方法，只使用了有限数量的基因来证明 mRNA 扩增的线性和重复性（Van Gelder et al. 1990；Eberwine et al. 1992；Lockhart et al. 1996；Phillips and Eberwine 1996；Kacharmina et al. 1999；Luo et al. 1999）。

然而，通常的 aRNA 扩增仍然容易受到偏差的影响，这是因为对 mRNA 5′端序列的代表性不足，以及由于在低温下进行双链 cDNA 的合成（Phillips and Eberwine 1996）会导致第二链合成中产生错误的引物引导。尽管不能充分代表 5′端序列在理论上可以通过使用发夹环引导 cDNA 第二链的合成来克服，与更为有效的Gubler-Hoffman方法（Kacharmina et al. 1999）相比，由于在低温下可能的非特异性引物延伸，因此仍旧不能保证没有偏离。所以，直到现在，仍然不清楚通常的 aRNA 扩增是否导致 mRNA丰度的显著偏离（Bertucci et al. 1999）。

在本方案中，如图 3-20 所示，描述了一个修改的方法来产生大量全长的 aRNA，用于高通量的 cDNA 微阵列分析。在通常的 aRNA 扩增方法（Eberwine et al. 1992；Phillips and Eberwine 1996；Kacharmina et al. 1999）的基础上，修改的方法中采用了模板转换，用反转录酶在 mRNA 模板的 5′端产生全长的双链 cDNA（Chenchik et al. 1998；Matz et al. 1999），可以使用 T7 RNA 聚合酶在体外转录。



205

图 3-20　使用模板转换扩增 mRNA

示意图中显示了将 mRNA 转化为双链 cDNA，可用于在噬菌体编码的 RNA 聚合酶作用下转录成 aRNA 或者是通过 PCR 扩增。方案中的步骤总结在下表中。关于进一步的详细情况，见本方案。（经许可修改，来自 CLONTECH Atlas Smart™探针扩增试剂盒用户手册© 2000 的图 1）

步　骤	模　板	引　物	酶	产　物
a	mRNA	T7-oligo(dT)	反转录酶	mRNA-cDNA 杂合子
b	mRNA-cDNA 杂合子	无	反转录酶	C-加尾的 mRNA-cDNA 杂合子
c/d	C-加尾的 mRNA-cDNA 杂合子	TS 引物	反转录酶，RNase H DNA 聚合酶	双链 cDNA
e	双链 cDNA	无	T7 RNA 聚合酶	aRNA

206　cDNA 第一链的合成是用 5′端含有能被噬菌体编码的 RNA 聚合酶识别的 T7 启动

子的 oligo(dT) -T7 引物启动的。cDNA 第二链的合成利用了 Mo-MLV 反转录酶内在的末端转移酶活性和模板转换能力（Chenchik et al. 1998；Matz et al. 1999；SMART PCR cDNA 合成试剂盒用户手册）。该酶在到达 mRNA 模板的 5′端时以非模板依赖方式向 cDNA 第一链的 3′端加上数个核苷酸残基（通常是 C 残基）。当 3′端带有一段短的加长的 dG 残基的寡核苷酸（TS 引物）加入到反应混合物中之后，与结合在新合成的 cDNA5′端突出的 dC 残基杂交。随后，反转录酶可以用 TS 引物作为模板，在 dC 区域的较远处延伸 cDNA，生成一个短片段的双链 DNA。用 RNase H 处理除去原始的 mRNA模板和加入热稳定的 DNA 聚合酶后，以 cDNA 第一链作为模板，TS 引物启动双链 cDNA 的合成。合成的全长双链 DNA 经过纯化后，用作在噬菌体编码的 RNA 聚合酶催化下进行体外转录的模板。产生的 aRNA 可以作为随机引物法标记 cDNA 的模板，或者是更多轮 PCR 扩增的模板。

该方法与链的长度无关，从而克服了使用末端转移酶催化向 cDNA 第一链的 3′端添加核苷酸时的偏差。线性扩增的效率在第一轮中大于 2000 倍，在第二轮中超过 800 000倍，并且可靠地反映了在初始的总 RNA 中观察到的基因表达的相对水平。最近通过基因表达谱分析对该方法的可靠性进行了评价（Wang et al. 2000），近来还通过应用实时定量 PCR 对扩增和未扩增的 RNA 中估计的基因表达水平进行比较（Ohnmacht et al. 2001）。这些实验证明在不同的 RNA 样品中基因表达的比例在多轮扩增后得到了保持。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

重要事项：本方案中所有试剂使用经 DEPC 处理的 H_2O 配制（见 **RNA 操作的特别注意事项**信息栏）。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

乙酸铵（7.5 mol/L）

二硫苏糖醇（DTT）（0.1 mol/L）〈!〉

将 0.31 g DTT 溶解于 20 ml 0.1 mol/L 醋酸钠（pH 5.2）。将溶液过滤灭菌并分装成小份保存在－20℃。在这些条件下，DTT 对空气氧化是稳定的。

乙醇（无水）

据报道，有些批次的乙醇（100%或 200 标准强度）中含有会使微阵列试验中产生很深荧光背景的污染物质，95%的乙醇不存在这个问题。为避免这个问题，可以用 95%的乙醇或已鉴定证明不会产生高荧光背景的无水乙醇。Warner-Graham 公司（Cockeysville, Maryland）出售的无水乙醇（USP 级乙醇，64-17-5）效果很好。

线性聚丙烯酰胺

线性聚丙烯酰胺是用乙醇沉淀皮克级的核酸时有效的中性载体（Gaillard and Strauss 1990）。要制备线性聚丙烯酰胺，在 40 mmol/L Tris-Cl、20 mmol/L 醋酸钠、1 mmol/L

EDTA(pH 8.0) 中配制5%的丙烯酰胺溶液（不含甲叉双丙烯酰胺）。加入1/100体积新鲜配制的10%过硫酸铵和1/1000体积的TEMED。当溶液变得黏稠时，用2.5倍体积的乙醇沉淀聚合物。振荡过夜将团块溶解于20倍体积的 H_2O。4℃下，0.25%的线性聚丙烯酰胺溶液可以稳定保存数年。

单相裂解溶液（TRIzol 15596，GIBCO）

NaOH(1 mol/L)，含有2 mmol/L EDTA

酚：氯仿：异戊醇（25：24：1)〈!〉

酶和缓冲液

Advantage DNA polymerase(CLONTECH) 和 Advantage PCR 缓冲液

本产品是一种氨基末端缺失的 *Taq* DNA 聚合酶和另一种具有3′→5′校正活性的热稳定DNA聚合酶的混合物。这两种酶的组合使得全长cDNA扩增更加有效，而且比常规的PCR错误率要低（Barnes 1994)。

cDNA 第一链缓冲液（由酶的生产商提供）

如果没有生产商提供的缓冲液，使用下面的5×第一链缓冲液：

250 mmol/L Tris-Cl(pH 8.3)

375 mmol/L KCl

15 mmol/L $MgCl_2$

体外转录试剂盒（Ambion T7 MEGAscript 试剂盒）

反转录酶

重组鼠反转录酶（RT）可以通过许多制造商得到。我们强烈推荐使用没有RNase H活性的突变形式。在本方案中使用的是SUPERSCRIPT Ⅱ反转录酶（Invitrogen/GIBCO)。然而，本方案很容易修改以适用于其他商品化的 RNase H^- 形式的Mo-MLV RT。反转录酶的特异活性取决于用于在 *E. coli* 中表达酶的特定cDNA克隆并且各个制造商之间存在差异。按照常用的各个制造商的说明使用酶（如单位/μg加入的RNA和温育温度)。Mo-MLV RT 对温度敏感，必须保存在－20℃直到需要时。

RNase H(2 单位/μl)

RNA 酶抑制剂

将抑制剂储存在－20℃。RNase的蛋白抑制剂可以与大多数RNase结合并抑制其活性，但是不影响野生型Mo-MLV RT的活性。许多制造商销售这些抑制剂，商品名也不相同（如Promega 的 Rnasin 和 5-Prime→3-Prime 的 Prime Inhibitor)。

核酸与寡核苷酸

ATP、GTP、CTP、UTP（每种浓度为75 mmol/L；Ambion MEGAscript T7试剂盒的组分）

dNTP/低浓度-dTTP 溶液

配制含有dATP、dCTP、dGTP的溶液，浓度均为5 mmol/L，且dTTP为2 mmol/L。将溶液分装成小份于－20℃保存。关于dNTP的配置和保存，见附录2。

dNTP 溶液，含有全部四种dNTP，每种浓度均为10 mmol/L

将溶液分装成小份于－20℃保存。关于dNTP的配置和保存，见附录2。

Oligo(dT)$_{15}$-T7 启动子

5′ AAA CGA CGG CCA GTG AAT TGT AAT ACG ACT CAC TAT AGG CGC(T)$_{15}$

该启动子序列比由 Dunn 和 Studier(1983) 定义的启动子的公认序列要长得多，New England Biolabs 和 Stratagene 有售。在这里显示的扩展序列中，公认序列嵌入在两侧翼区序列之间，在其 5′侧翼区域，为 T7 RNA 聚合酶结合提供了空间，3′侧翼区域的三个核苷酸可以增加 T7 的酶转录活性。

合成的寡核苷酸要使用聚丙烯酰胺凝胶电泳进行纯化（Sambrook and Russell 2001）并以约 25 pmol/μl（约 0.5 μg/μl）的浓度溶解于 H_2O 中。

随机引物

由于具有一致的长度，且没有序列偏向性，所以随机序列的合成寡核苷酸是有效地将 aRNA拷贝为 cDNA（本方案）首选的引物。最佳长度的寡核苷酸（六聚体、七聚体或八聚体；Suganuma and Gupta 1995）可以通过商业途径购买（如 Pharmacia 或 Boehringer Mannheim）或者就地在自动 DNA 合成仪上合成。将引物按照浓度为 1.0 μg/ml 溶解在 TE (pH 7.6) 中。引物溶液分装成小份于−20℃下保存。

总 RNA

本方案需要的总 RNA 量为 0.01～5 μg，按照方案 1 中介绍的方法制备。哺乳动物细胞的总 RNA 的含量为 20 ～40 pg(Roozemond 1976；Ueniura 1980)，并且其中只有 0.5～1.0 pg 是 mRNA。本方案经过优化，适合于用本部分前面描述的方法（请见第 3 章 Part Ⅰ 的方案）对纯化自合适来源的 500～2.5×10^5 个细胞进行抽提得到的 RNA。

TS 引物 5′ AAG CAG TGG TAA CAA CGC AGA GTA CGC GGG 3′

根据 Chenchik et al(1998) 的数据，当 3′GGG 序列使用核糖核苷酸而非脱氧核糖核苷酸合成时得到最佳的结果。然而，作者多次使用含有…dGdGdG3′的 TS 引物，并没有任何问题。合成的引物必须经过电泳纯化（Sambrook and Russell 2001）并按照约 30 pmol/μl（约 0.5 μg/μl）的浓度溶解于水中。

离心机和转子

台式离心机（高速 16 000 *g*）用于 1.5 ml 微量离心管

专用设备

Micro Bio-Spin 柱（Bio-Gel P6，Bio-Rad）

PCR 管（0.2 ml，薄壁管）

Phase Lock Gel(heavy)(0.5 ml 试管；Brinkmann Eppendorf)

热循环仪

42℃和 65℃水浴

方法

合成 cDNA 第一链和第二链

1. 为合成 cDNA 第一链，在 0.2 ml 薄壁 PCR 管中混合以下试剂：

总 RNA	0.01～0.5 μg

oligo(dT)$_{15}$-T7 启动子（100 pmol/ml）　　1 μl

H_2O　　加至 8 μl

将试管密封后在 70℃下温育 5 min，从水浴或热循环仪中取出试管并使其在室温下停留 5 min。

oligo(dT) -T7 引物的用量随反应中使用的总 RNA 的量而变化。虽然引物的摩尔量总应该是过剩的，但是太多的引物会导致本方案后面的阶段发生非特异性的扩增（Baugh et al. 2001）。在使用珍贵的样品之前，要进行一系列包含不同量的 RNA 模板和不同摩尔比引物的试验反应和扩增，目的是寻找使 aRNA 产量最高的引物：模板的比例。

2. 向RNA/oligo(dT) 混合物中加入：

5×第一链缓冲液	4 μl
TS（模板转换）引物（30 pmol/μl）	1 μl
0.1 mol/L DTT	2 μl
RNase 抑制剂	1 μl
10 mmol/L dNTP 溶液	2 μl
SUPERSCRIPT Ⅱ RT(200 单位/μl)	2 μl

轻轻涡旋混合试剂并将反应混合物短时离心去除气泡。将密封的微量离心管放在热循环仪中，42℃下温育 90 min。

3. 为合成 cDNA 第二链，直接向第一链反应混合物中加入下列试剂：

H_2O	106 μl
Advantage PCR 缓冲液	15 μl
10 mmol/L dNTP 溶液	3 μl
RNase H（2 单位/μl）	1 μl
Advantage cDNA 聚合酶混合物（CLONTECH）	3 μl

轻轻涡旋混合试剂并将试管放入热循环仪中，按照以下步骤设定程序：

37℃ 5 min（消化 cDNA 第一链杂合体中的 mRNA 组分）

94℃ 2 min（使 cDNA 第一链 3′端的 DNA-DNA 杂合体变性）

65℃ 1 min（使引物发生结合）

75℃ 30 min（使引物延伸成全长的 cDNA 链）

由于启动 cDNA 第二链合成的 TS 引物已经存在于 cDNA 第一链合成的反应混合物中，并且被用作产生向 cDNA 第一链 3′端延伸的引物，因此本步骤不需要再加入引物。

210

4. 加入 7.5 μl 含有 2 mmol/L EDTA 的 1 mol/L NaOH 溶液，并将反应在 65℃下温育 10 min 来终止反应。

NaOH 升高了反应混合物的 pH，增加了酚：氯仿：异戊醇的抽提效率（步骤 5）并使酶失活。到这一步时，反应可在－20℃下储存。

纯化双链 cDNA

本步骤用于防止未掺入的 dNTP、引物和失活的酶被带入后续的体外转录反应。

5. 将反应混合物转移到含有以下试剂的 1.5 ml 微量离心管中：

线性聚丙烯酰胺（0.1 μg/μl）　　1 μl

酚：氯仿：异戊醇　　150 μl

密封试管并轻轻涡旋混匀试剂。

6. 将乳液转移到 Phase Lock Gel 试管中，并将试管在室温下以 10 000 *g* 离心 5 min。
7. 将水相转移到新的 1.7 ml 离心管中并加入 70 μl 7.5 mol/L 乙酸铵，然后加入1 ml 无水乙醇。密封试管并轻轻涡旋混匀试剂。不作停留，立即进行下一步操作。
8. 在室温下，将乙醇溶液以 10 000 *g* 离心 20 min（以避免寡核苷酸引物共沉淀）。
9. 弃去上清液并用 800 μl 75％乙醇洗涤团块。再次以 10 000 *g* 离心 10 min。重复洗涤步骤一次。
10. 弃去最后的上清液并让团块在空气中干燥。将双链 cDNA 溶解在 70 μl H_2O 中。
11. 按照下面的方法准备 Micro Bio-Spin 层析柱（Bio-Gel P6 柱，在 Tris 中）：
    a. 将层析柱反转数次使沉积的凝胶重悬并除去气泡。
    b. 折断尖头并将层析柱放入一个收集管中。
    c. 移去管盖，让缓冲液流净。
    d. 弃去流出的缓冲液，并向柱中加入 700 μl H_2O。
    e. 弃去流出液体，并重复洗涤步骤 3 次。
    f. 室温下将层析柱以 700 *g* 离心 2 min。

    打开柱子时，凝胶可能会粘在管盖下壁。尽可能多的回收凝胶并放回柱中，然后吸掉盖上残留的任何凝胶。

    由于制造商提供的收集管不是无 RNase 活性的，需使用自备的无 RNase 活性试管来收集洗脱液体。
12. 将层析柱放入新的无 RNase 活性的收集管中。向柱中央加入制备的 DNA（70 μl，步骤 9）并将层析柱以 700 *g* 离心 4 min。
13. 在旋转蒸发器（如 SpeedVac）中将样品干燥至半干，并溶解在 8 μl H_2O 中。

211

体外转录双链 cDNA

所有试剂与储存液（T7 RNA 聚合酶除外！）在使用之前都要完全解冻，混匀，并短时离心。将聚合酶保存在冰箱中，直到主要混合物的其他试剂已经全部加入。检查 MEGAscript 试剂盒中的 10×反应缓冲液中是否有沉淀。如果有可见沉淀，将缓冲液加热至 37℃直到沉淀消失（5～10 min）。轻轻涡旋缓冲液并短时离心，使液体集中在试管底部。

方案中本部分使用的试剂由 Ambion T7 MEGAscript 试剂盒提供。

14. 向双链 cDNA 样品（步骤 13）中加入：

    ATP（75 mmol/L）　　2.0 μl

    GTP（75 mmol/L）　　2.0 μl

    ATP（75 mmol/L）　　2.0 μl

    GTP（75 mmol/L）　　2.0 μl

    10×反应缓冲液　　2 μl

    酶混合物（RNase 抑制剂和 T7 RNA 聚合酶）　　2 μl

将反应混合物在37℃下温育6h。

根据Ambion，温育过程可以中断并将反应试管在－20℃下保存，以后再继续温育不会影响效率。

纯化 aRNA

15. 使用标准的酚：氯仿抽提（方案15中步骤6所述）来纯化步骤14中合成的RNA。也可使用单相试剂如TRIzol，简述如下：
    a. 向转录反应中加入1 ml TRIzol，用移液管或通过轻轻涡旋混匀试剂。
    b. 向每1 ml TRIzol溶液中加入200 μl氯仿，倒转试管15 s混匀试剂，将试管在室温下竖立2～3 min。
    c. 在4℃下将试管以10 000 *g* 离心15 min。
    d. 将水相转移到新的试管中，然后向每1 ml TRIzol试剂加入500 μl异丙醇。
    e. 将样品在室温下静置10 min，然后以10 000 *g* 离心15 min。
    f. 用1 ml 70%乙醇洗涤团块，重复一次。
    g. 让团块在空气中干燥15～20 min，然后溶解在20 μl H_2O中。
    h. 使用比色法测量RNA的浓度。

212

第二轮扩增

16. 在0.2 ml薄壁PCR管内混合：

aRNA（步骤15）	0.5～1 μg
随机引物（1 μg/μl）	2 μl
H_2O	加至10 μl

将混合物加热到70℃并保持5 min，然后在工作台上冷却5～10 min。

17. 向步骤14制备的混合物中加入下列试剂：

5×第一链缓冲液	4 μl
oligo(dT)$_{15}$-T7引物（100 pmol/ml）	1 μl
0.1 mmol/L DTT	2 μl
RNase抑制剂	1 μl
10 mmol/L dNTP溶液	2 μl
SUPERSCRIPT Ⅱ RT（200单位/μl）	2 μl

轻轻涡旋混合试剂并将反应混合物短时离心去除气泡。将密封的微量离心管放在热循环仪中，42℃下温育90 min。

18. 从此处开始，按照前面所述的cDNA第二链的合成方法（步骤3～4），纯化双链cDNA（步骤5～13），用40 μl反应混合物替代20 μl进行体外转录（步骤14），并分离aRNA（步骤15）。

第二轮aRNA的品质可以通过变性胶电泳进行检测（Sambrook and Russell 2001）。第一轮扩增获得的aRNA大小为500～6000个核苷酸，第二轮扩增为500～4000个核苷酸（见表3-8和图3-21）。

此时aRNA能用作合成荧光标记cDNA的模板，由方案13中所述的反转录酶催化。

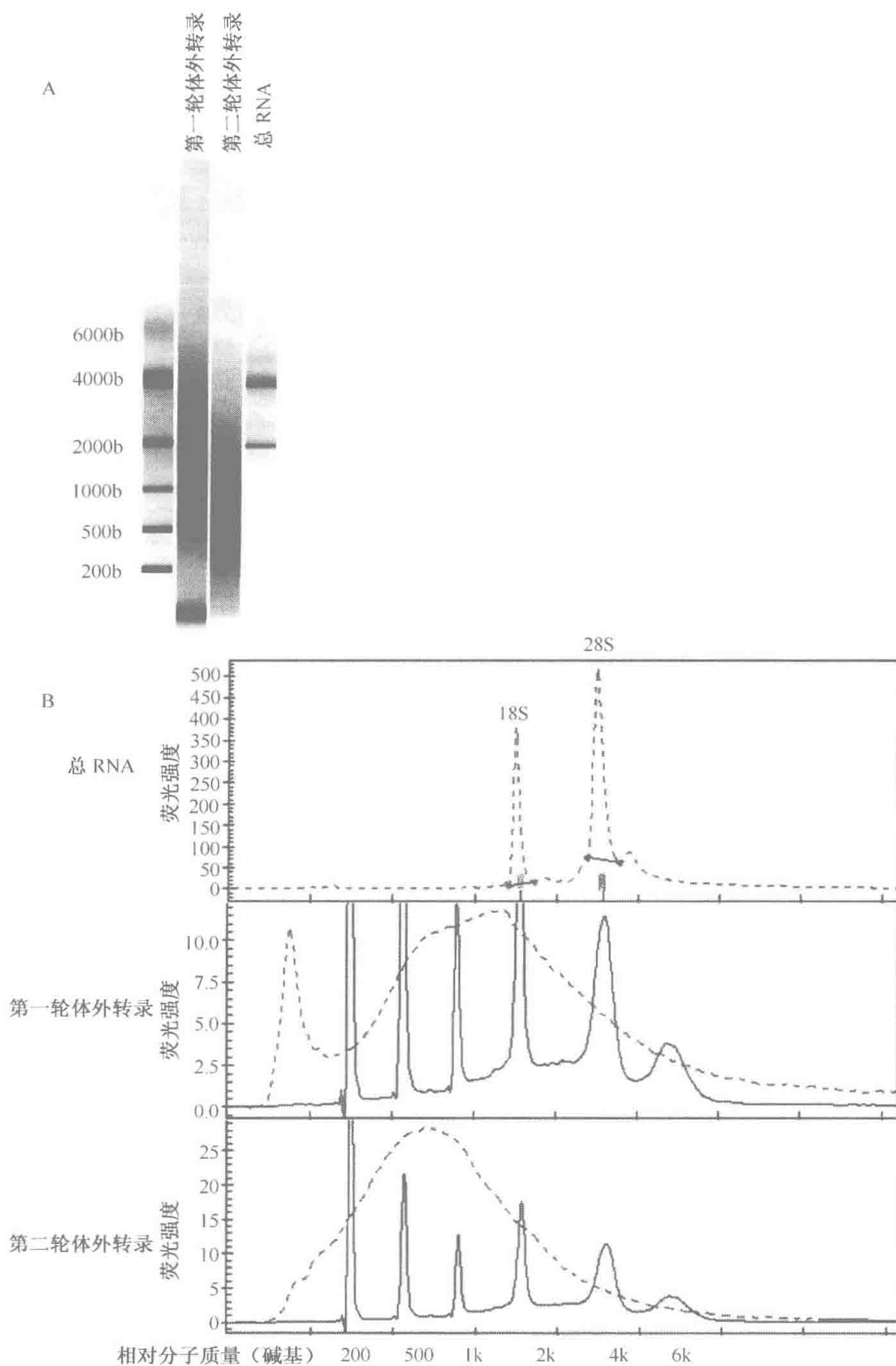


图 3-21　体外转录（IVT）产物分析

A. 使用 Agilent Bio-Analyzer 2100 在 RNA 6000 芯片上通过电泳分离体外转录的 RNA 产物；B. 使用 Agilent Bio-Analyzer 得到的扩增 RNA 的质量分析电泳图谱。（实线）分子质量，标准（虚线）样品。

表 3-8 RNA 的浓度和比值

样品 Mel 888	总 RNA
片段名称	
1	**18S**
2	**28S**
RNA 修正面积	2386.00
RNA 浓度/(ng/μl)	1477.54
rRNA 比值(28S/18S)	**1.58**
	第一轮 IVT
RNA 修正面积	529.66
RNA 浓度/(ng/μl)	327.99
rRNA 比值(#2/#1)	**0**
	第二轮 IVT
RNA 修正面积	1057.73
RNA 浓度/(ng/μl)	655
rRNA 比值(#**2**/#**1**)	**0**

数据来自图 3-21B。

方案 17

用 *E. coli* DNA 聚合酶的 Klenow 片段标记第二链 DNA

在下面的方案中，通过随机六聚体引导和 *E. coli* DNA 聚合酶 Klenow 片段催化的反应使用荧光标记的核苷酸标记制备的单链 cDNA。由于合成的随机序列的寡核苷酸具有一致的长度且无序列偏向性，因而是标记 cDNA 首选的引物，如从 RNA 扩增获得的 cDNA（方案 15 和方案 16）。在本方案结尾的替代方案中描述了使用 Klenow 片段标记双链基因组 DNA。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

重要事项：所有缓冲液和溶液使用经 DEPC 处理的 H_2O 配制并分装成小份冰冻保存。

缓冲液和溶液

EDTA（0.5 mol/L，pH 8.0）

酶和缓冲液

E. coli DNA 聚合酶Ⅰ的 Klenow 片段（如来自 New England Biolabs 的高浓度的、无核酸外切酶活性的 Klenow，50 单位/μl）

5× 随机引物缓冲液

250 mmol/L Tris(pH 8.0)

100 mmol/L NaCl

25 mmol/L $MgCl_2$

10 mmol/L 二硫苏糖醇（DTT）〈!〉

1 mol/L HEPES（用 4 mol/L NaOH〈!〉将 pH 调至 6.6）

将缓冲液过滤灭菌，并分成小份储存于－20℃。

RNase H(1～5 单位/μl)

核酸

Cy3 标记的 dUTP(1 mmol/L ）和 Cy5 标记的 dUTP(1 mmol/L ）（Amersham Biosciences)

结合了羰花菁染料的 dNTP 对光敏感，应该于－20℃下保存在避光的容器中。

10×dNTP/低浓度-dTTP 溶液

配制含有 dATP、dCTP、dGTP 的溶液，浓度均为 5 mmol/L ，且 dTTP 为 2 mmol/L 。将溶液分装成小份于－20℃保存。关于 dNTP 的配置和保存，见附录 2。

$Oligo(dT)_{12\text{-}18}$或者锚定引物（2 μg/μl）

215

随机引物（4 μg/μl 溶于 pH 7.6 的 TE 中）

最佳长度的寡核苷酸（六聚体、七聚体或八聚体；Suganuma and Gupta 1995）可以通过商业途径购买（如 Pharmacia 或 Boehringer Mannheim）或者就地在自动 DNA 合成仪上合成。引物溶液分装成小份于－70℃下保存。

总 RNA 或 $poly(A)^+$ RNA

按照方案 1、方案 5 或者方案 6 制备。

离心机和转子

台式离心机（高速 16 000 g）用于 1.5 ml 微量离心管

专用设备

Microcon YM-100(42412，Amicon)

Molecular Dynamics Storm 荧光扫描仪（或与之相当的）

PCR 管（0.2 ml，薄壁管）

QIAquick PCR 纯化试剂盒（50）(28104，QIAGEN)

热循环仪

循环仪要能容纳 0.2 ml 薄壁 PCR 管。

方法

mRNA 和引物复性

1. 在薄壁 0.2 ml PCR 管中，混合：

总 RNA	50 μg
或	

poly(A)$^+$ RNA	1 μg
oligo(dT)$_{12\text{-}18}$或锚定引物（2 μg/μl）	0.5 μl
H_2O	加至 22.5 μl

2. 将反应混合物在热循环仪中于 65℃下加热 10 min，之后在冰上冷却 2 min。

未标记第一链 cDNA 的合成

3. 向来自步骤 2 的混合物中加入：

5×第一链 cDNA 缓冲液	8.0 μl
0.1 mol/L DTT	4.0 μl
5 mmol/L dNTP	4.0 μl
SUPERSCRIPT Ⅱ RT(200 单位/μl)	1.5 μl

反转录酶对在气/液界面的变性非常敏感，因此当混合反应的组分时，要非常小心避免产生泡沫。这有时很难做到，因为许多反转录酶制备物的缓冲液中含有非离子型去污剂，如 Nonidet P-40 或者 Triton X-100。

Mo-MLV（SUPERSCRIPT Ⅱ）对温度敏感，应该储存在－20℃下直到使用前的最后时刻，并且在第一时间放回冰箱。

216

4. 将反应混合物在 42℃下温育 1 h。
5. 加入 0.25 μl RNase H(1～5 单位/μl)。
6. 将反应混合物在 37℃下温育 30 min。

纯化 cDNA 第一链

7. 加入 160 μl TE(pH 7.4) 并轻轻涡旋充分混合反应混合物。
8. 将反应混合物转移到 Microcon YM-30 超滤管中，以 16 000 *g* 离心 5 min，弃去流出液体。
9. 向超滤管中加入 200 μl TE(pH 7.4)。再次以 16 000 *g* 离心 5 min，弃去流出液体。
10. 在新的收集管中倒转超滤管以回收浓缩物。将倒转的超滤管以大约 1000 *g* 离心 1 min。浓缩物的体积应当约为 2 μl。

    在通过离心过滤浓缩溶液时，达到希望的最终体积所需要的离心时间是变化的。过长时间的离心会将溶液中的溶剂几乎全部去除。荧光标记的核酸在过滤器上浓缩成这种方式的时候，它们的回收非常困难。为了减少这一问题，用保守估计的离心时间逐渐接近希望的体积。

11. 向超滤管中加入 40 μl H_2O。将倒转的柱子放回同一支收集管中（步骤 10）。将倒转的超滤管以大约 1000 *g* 离心 1 min。
12. 用 H_2O 将流出液体的体积调整到 58 μl。
13. 将 29 μl 流出液体转移到新的 0.2 ml 薄壁 PCR 管中，剩余的 29 μl 可以保存于－20℃或者用于标记。

荧光标记 cDNA 第二链的合成

14. 混合下列各组分：

cDNA 第一链（步骤 13）	29 μl
10×Klenow 缓冲液	4 μl
随机引物（4 μg/μl）	1 μl

15. 将试管置于热循环仪中在 100℃下温育 2～3 min，在工作台上冷却至室温。
16. 向试管中加入下列组分：

10×dNTP/低浓度-dTTP 溶液	4 μl
Cy3 标记的或者 Cy5 标记的 dUTP（1 mmol/L）	1 μl
Klenow 片段（5 单位/μl）	1 μl

将反应在 37℃下温育 3 h。

纯化荧光标记的 cDNA 第二链

17. 使用 QIAquick PCR 纯化柱（QIAGEN）从荧光标记的 DNA 中除去污染物。根据方案 14 中步骤 12 下方信息栏中的说明进行操作。

替代方案：用 *E. coli* DNA 聚合酶的 Klenow 片段标记基因组 DNA

在本方案中，通过 *E. coli* DNA 聚合酶 Klenow 片段的催化和随机六聚体引导的反应给片段化的基因组 DNA 标记上荧光核苷酸。标记的 DNA 可用作比较基因组杂交实验的靶分子（见第 5 章）。

附加材料

基因组 DNA，使用超声粉碎为平均长度 500 bp 至 1 kb 的片段。

按照 Sambrook 和 Russell(2001) 在 12.15--12.16 所描述的方法制备超声粉碎的 DNA。

Microcon YM-30(Millipore)

5×三种 dNTP 溶液，包含 dATP、dCTP、dGTP，每种浓度均为 5 mmol/L

水浴，设定在 95℃ 和 37℃

方法

1. 在1.5 ml 的微量离心管中，将片段化的基因组 DNA 和随机引物混合：

片段化的基因组 DNA	1 μg
随机引物	2 μl
H_2O	加至 20 μl

2. 在热循环仪中将反应混合物加热到 95℃下 2 min，然后在冰上冷却 2 min。
3. 准备标记混合物：

5×随机引物缓冲液	10 μl
5×三种 dNTP 溶液	5 μl
H_2O	12 μl

4. 将标记混合物转移到含有基因组 DNA 的试管中，并加入：

Cy3 或 Cy5 标记的 dUTP(1 mmol/L)	2 μl

Klenow 片段（50 单位/μl）　　　　1 μl

将反应在暗处 37℃下温育 3 h。

5. 加入 2.5 μl EDTA（0.5 mol/L，pH 8.0）来终止反应。
6. 按照下面的方法除去未掺入的 dNTP：
   a. 向反应中加入 470 μl H_2O。
   b. 将稀释的反应混合物转移到放在收集管中的 Microcon YM-30 过滤装置内。
   c. 以 13 000 *g* 离心 8～10 min，使体积减少到 20～30 μl。
   d. 向 Microcon YM-30 上部的 20～30 μl 液体中加入 450 μl 的 H_2O。
   e. 重复离心步骤。
   f. 向 Microcon YM-30 上部的 20～30 μl 液体中加入 450 μl 的 H_2O。
   g. 重复离心步骤直到体积减少到大约 10 μl。
   h. 将过滤装置放入新的微量离心管中并以 3000 *g* 离心 3 min。

   将标记的基因组 DNA 放在暗处于 4℃下保存直到用于杂交。

218

方案 18

用生物素化核苷标记扩增 RNA 来与寡核苷酸阵列杂交

Ed Tom，Jerry Lee，Dorian Willhite，Suzanne Ybarra，Richard Glynne（Eos 生物技术公司，南旧金山，加利福尼亚，94080）

一般来说，制备的 mRNA 不用于直接同寡核苷酸微阵列杂交，取而代之的是合成的含有标记核苷酸的 mRNA 的衍生物。在本方案中，mRNA 转化为带有噬菌体 T7 启动子的未标记双链 cDNA。然后 cDNA 用作在生物素化核苷酸存在下，由 T7 DNA 依赖的 RNA 聚合酶催化进行体外转录的模板。使用生物素化的碱基进行标记，是因为荧光标记的核苷酸是噬菌体编码的 RNA 聚合酶的较差底物。由于每个 cDNA 模板被转录许多次，mRNA 序列（或者更准确地说是与它们互补的拷贝）以线性方式扩增了几千倍，而且，生物素化的碱基标记用于寡核苷酸阵列上杂交体的检测。生物素化的 cRNA 通过化学方法降解成片段，与对照 cRNA 片段混合，然后与寡核苷酸阵列杂交。生物素化的 cRNA 通过与链亲和素——藻红蛋白的结合进行检测，并使用共焦激光扫描仪进行扫描。

本标记方案已经用点样的寡核苷酸阵列进行测试，但还没有用于 cDNA 阵列。结果显示，寡核苷酸-生物素化 RNA 杂交体可以和用 Cy3 或 Cy5 标记的杂交体以几乎相等的灵敏度被检测到。

带有 T7 启动子的双链 cDNA 的转录，最初是用来扩增少量的 mRNA 而不会显著影响 mRNA 的拷贝水平（Van Gelder et al. 1990；Eberwine et al. 1992；Phillips and Eberwine 1996；Luo et al. 1999；Wang et al. 2000）。该方法的第一阶段——将 mRNA 转化为带有 T7 的双链 cDNA——在方案 15 的步骤 1～4 中给出了详细的说明，在开始这里描述的标记方案之前应当完成。

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

重要事项：本方案中所有试剂使用经 DEPC 处理的 H_2O 配制（见 **RNA 操作的特别注意事项**信息栏）。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

乙酸铵（5 mol/L）

乙醇（无水）置于室温下，和 80%并冷却到−20℃

乙醇（无水）

据报道，有些批次的乙醇（100%或 200 标准强度）中含有会使微阵列试验中产生很深荧光背景的污染物质，95%的乙醇不存在这个问题。为避免这个难题，可以使用 95%的乙醇或者已鉴定的性能可靠的品牌无水乙醇。购自 Rossville(Gold Shield) 的无水乙醇（USP 级乙醇）效果一致性很好。

酚：氯仿：异戊醇（25：24：1)〈!〉

酶和缓冲液

T7 RNA 聚合酶（Ambion MEGAscript T7 试剂盒的成分）

10×T7 转录缓冲液（Ambion MEGAscript T7 试剂盒的成分）

核酸和寡核苷酸

ATP、GTP、CTP、UTP（每种浓度均为 75 mmol/L；Ambion MEGAscript T7 试剂盒的成分）

Bio-11-CTP(10 mmol/L)（Enzo）

Bio-16-UTP(10 mmol/L)（Roche）

双链 cDNA，带有噬菌体 T7 启动子序列

按照方案 15，步骤 1～4 所述的方法从 10 μg 总 RNA 或者 1 μg poly(A)$^+$ RNA 制备的 cDNA；cDNA 合成后的体积应为约 155 μl。

DNA 或 RNA 分子质量标记/序列梯（MBI Fermentas）

专用试剂和试剂盒

线性聚丙烯酰胺（0.25%）（见方案 16）

RNeasy 试剂盒（QIAGEN）

T7 MEGAscript 试剂盒（Ambion）

专用设备

PCR 管（0.2 ml，薄壁的）

对于大量的样品，可以使用排管和管盖。

Phase Lock Gel Light 试管（Brinkmann Eppendorf 或者 5 Prime→3 Prime）

Picofuge(Stratagene) 或者与之相当的离心机

带加热盖的热循环仪（MJ Research）或者与之相当的

方法

浓缩 cDNA

将总 RNA 或 poly(A)$^+$RNA 转化为带有噬菌体 T7 启动子的双链 cDNA 后（步骤 1～4，方案 15）后，按照下述方法在 Phase Lock Gel Light 试管中纯化 cDNA。

1. 为每个反应将一支 Phase Lock Gel Light 试管在室温下以 13 000 g 或更高的转速离心 1 min。
2. 向 Phase Lock Gel Light 试管中加入 cDNA 样品（约 155 μl）。
3. 加入等体积（约 155 μl）的酚：氯仿：异戊醇（25：24：1）。
4. 倒转试管并剧烈振荡或者涡旋 1～2s。
5. 将 Phase Lock Gel Light 试管在室温下以 13 000 g 或更高的转速离心 2 min。
6. 将上部的水层（约 150 μl）转移到新的微量离心管中。
7. 加入下列试剂沉淀 cDNA：

线性聚丙烯酰胺（0.25%）	0.5 μl
水相	150 μl
5 mol/L 乙酸铵（与水相体积相等）	150 μl

   轻轻涡旋混匀试剂。
8. 加入 2.5×水相体积的乙醇和 5 mol/L 乙酸铵，通常为 750 μl。轻轻涡旋混合试剂，然后将溶液在－20℃下至少静置 20 min。
9. 在 16℃下以 12 000 g 或更高的转速离心 30 min 来回收 cDNA。
10. 仔细除去暗白色团块以外的乙醇。
11. 用 500 μl 预冷到－20℃的 80%乙醇洗涤团块两次，每次在 16℃下以 13 000 g 或更高的转速离心 5 min 来回收 cDNA。
12. 离心后，小心除去乙醇并将开口的离心管放在工作台上或者化学通风橱内，直到团块半干。
13. 将团块溶解在 3 μl H_2O 中。

使用 Ambion T7 MEGAscript 试剂盒，在体外转录 cDNA

14. 准备转录反应主要混合物。将每个反应的体积乘以（x+1/2），其中 x 是全部反应的总数。

ATP(75 mmol/L)	2.0 μl
GTP(75 mmol/L)	2.0 μl
CTP(75 mmol/L)	1.5 μl

UTP(75 mmol/L)	1.5 μl
Bio-11-CTP(10 mmol/L)	3.75 μl
Bio-16-UTP(10 mmol/L)	3.75 μl
10×T7 转录缓冲液	2.0 μl
T7 RNA 聚合酶（100 单位/μl）	2.0 μl

15. 向 0.2 ml 薄壁 PCR 管中加入 18.5 μl 的转录反应主要混合物。
16. 加入 1.5 μl 带有 T7 启动子的双链 cDNA（s）（步骤 13），轻轻吹吸混匀试剂，将未使用的 cDNA 保存于－20℃。
17. 在 picofuge 中将试管短时离心，使液体集中在试管底部。
18. 将试管放入热循环仪中并在 37℃下温育 6h，并将温育后的保持温度自动设定为 4℃。使用带有加热盖的热循环仪，以防止冷凝。

    生物素化的 RNA 可以在－20℃存放最多 1 周，或者在－80℃存放更长时间。预期的产量为从大约 0.5 μg 双链 cDNA 可以获得 50～100 μg 的 cRNA。

用 QIAGEN Rneasy 试剂盒纯化生物素化的 RNA

19. 向每份生物素化的 RNA 样品（步骤 18）中加入 80 μl 的 H_2O。轻轻吹吸混匀试剂，然后将混合物转移到新的 1.5 ml 微量离心管中。
20. 加入 350 μl RLT 缓冲液（QIAGEN），轻轻涡旋充分混合。
21. 加入 250 μl 无水乙醇，轻轻涡旋混合。
22. 将样品（700 μl）加入 RNeasy 离心柱中。
23. 将柱子在室温下用微量离心机以 10 000 *g* 离心 15～20 s。
24. 将流出液体从收集管转移到新的微量离心管中，保存在室温下，直到 RNA 的回收得到确认。
25. 弃去收集管。

    重要事项：用新的收集管替换它。

26. 向离心柱中加入 500 μl RPE 缓冲液（QIAGEN）。

    按照 QIAGEN 手册的规定，新开瓶的 RPE 缓冲液在使用前需要和 95%～100%乙醇混合。

27. 在室温下以 10 000 *g* 的转速将柱子离心 15～20 s。
28. 弃去流出液体，然后重复步骤 26～27。
29. 弃去流出液体，在室温下将柱子以 16 000 *g* 或更高的转速离心 2 min。

    离心结束后，检查确保在离心柱底部的尖头没有可见的缓冲液。如果有缓冲液存在，在吸去收集管中的流出液体后重复离心步骤。

30. 将离心柱转移至新的微量离心管中。向 RNeasy 膜的中央直接加入 50 μl 的 H_2O，并将柱子在室温下放置 5 min。
31. 在室温下将柱子以 16 000 *g* 或更高的转速离心 2 min。
32. 再次向 RNeasy 膜的中央直接加入 50 μl 的 H_2O，并将柱子在室温下放置 5 min。
33. 在室温下将柱子以 16 000 *g* 或更高的转速离心 2 min。
34. 弃去离心柱，用比色法在 260 nm 和 280 nm 处测定流出液体中 RNA 的量。

35. 使用1%的琼脂糖凝胶通过电泳分析200～300 ng的RNA。关于分子质量标记，可以使用RNA序列梯（小范围）或者DNA序列梯（间隔100 bp）。

验证RNA完整性时，天然的琼脂糖凝胶与完全变性的凝胶效果相同。相对于100 bp间隔的DNA序列梯，cRNA呈扩散状并且中间大小位于600～1000 bp。

加样到天然的琼脂糖凝胶上之前，cRNA不需要加热处理。

与寡核苷酸微阵列杂交之前，cRNA用下面的附加方案中所述方法进行片段化。

附加方案：生物素化cRNA的片段化

为了加快杂交动力学并减少二级结构，主要方案中制备的生物素化cRNA要在镁离子存在的条件下，加热处理成为50～100 bp的片段产物。

附加材料

生物素化cRNA
过滤器（0.2 μm）
乙酸镁
乙酸钾
Tris-乙酸（1 mol/L，pH 8.1）

Trizma碱。用乙酸调节pH。

方法

1. 将体积小于或等于16 μl的15 μg生物素化cRNA转移到PCR管中（如果需要，可使用乙醇沉淀进行浓缩）
2. 用无RNase酶的H_2O将cRNA的体积调整到16 μl。
3. 准备5×片段化缓冲液：

试剂	数量	最终浓度
1 mol/L Tris-乙酸(pH 8.1)	4.0 ml	200 mmol/L
乙酸镁	0.64 g	150 mmol/L
乙酸钾	0.98 g	500 mmol/L

用无RNase酶的H_2O将最终体积调整为20 ml，并用0.2 μm过滤器进行过滤。

4. 向cRNA中加入4 μl 5×片段化缓冲液，将混合物放在带有加热盖的热循环仪中于95℃下温育35 min。

重要事项：不推荐在加热装置中使用微量离心管，因为在温育过程中缓冲液会蒸发。

5. 将试管在picofuge中离心以收集冷凝液体并将试管放置在冰上。

片段化的cRNA可以用于杂交反应。

方案 19

未掺入的 Cy 染料的回收：使用 HPLC 重新纯化[①]

223

Chih Long Liu（生物学与生物医学科学中心，生物医学科学部，哈佛医学院，波士顿，马塞诸塞州，02115）

微阵列分析的一个限制是为 cDNA 微阵列杂交和分析所进行的 cDNA 标记中荧光核苷的花销太大。在典型的标记反应中，反转录酶将 mRNA 拷贝成 Cy3 或者 Cy5 标记的 cDNA，消耗约 3 nmol 的染料。按照目前的市场价格（295 美金/25 nmol，2001 年 10 月），如果需要分析大量样品，花销会很快变得令人不敢问津（一个典型的肿瘤表达谱分析可能包括数百个样品，仅在染料上的花费就需要数万美元）。

反转录酶不能有效地将结合了染料的核苷酸掺入到 cDNA 中。据报道，有高达 5% 的掺入率，而典型的掺入率约为 2%～4%（Yu et al，1994，未发表）。导致低掺入率的因素包括：①染料（尤其是 Cy5 dUTP）的不稳定性；②大的荧光生色团的空间阻碍。反应中使用的结合了染料的核苷酸的浓度一般接近于 K_m，能够生成满意产量的产物，并标记上合适的特定活性，能够有效地杂交。在标记反应中降低 Cy3 和 Cy5 标记核苷酸的浓度不会减少花费，因为已经很低的向 cDNA 中的掺入率会进一步地下降。

然而，通过从反应混合物中回收未掺入的染料可以降低费用。这里介绍的方法中，基于一种通过高效液相色谱（HPLC）分离 dNTP 的方案，从含有未掺入的未标记核苷酸、反转录产酶、缓冲液成分、降解的 RNA 和 DNA 引物以及丢弃的 cDNA 转录物使用过的反应混合物中回收染料。HPLC 柱含有阴离子交换的反相 C-18 树脂。通过阴离子交换从树脂上洗脱掉反应混合物中的阴离子和中性组分。由于柱子的反相性质，未掺入的未标记核苷酸因为没有花青染料修饰核苷酸而具有的碳氢链，具有短的保留时间。阴离子交换和反相两者结合的优点在于可以将希望得到的三磷酸盐形式的染料——核苷酸和单与双磷酸盐形式核苷酸分离开来。保留时间比花青修饰的核苷酸还长的最有可能是丢弃的 cDNA 转录物。请注意本方案不能用于回收氨基-烯丙基染料。

224

一般来说，经过纯化的回收 Cy3 和 Cy5 染料的活性与通常的染料是相当的，并且存在的活性差异是位于阵列与阵列之间/不同工作日之间的杂交差异范围之中。由于没有化学上的理由使得普通的染料和重新纯化的染料之间有很大的反应差异，因而任何观察到的假象很可能是纯化过程本身的问题或者由于某些原因被加入到原始标记反应中的不寻常的组分造成的。本方案包括：

- 为向柱子上大量进样而进行的最初的脱盐步骤。
- 冻干步骤以减少样品的体积。

---

① Kathleen Hayashibara(Perkins-Elmer Biosystems) 和 Linda McAllister(Celera Diagnostics)，Pat Brown 实验室先前的成员，协助了本方案的编写。通过 HPLC 纯化 Cy 染料的方法是根据 Herschlag 实验室前成员 Suzanne Admiraal 的方案。

- 用 Microcon 或者 Centricon 过滤器过滤除去原始标记反应中未掺入的物质或者高分子质量的组分。
- 通过 HPLC 纯化。
- 多轮冻干-重悬以蒸发掉 HPLC 纯化中使用的 TEAA 缓冲液。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

乙腈〈!〉

花菁染料核苷酸（Cy 标记的 dNTP），从标记反应中回收

花菁染料核苷酸的处理：从标记反应中收集 Cy 标记的 dNTP 通常在标记的 DNA 通过 Microcon过滤器浓缩阶段进行。在 Brown 实验室中，过滤器的流出液体分别收集在不同的试管中，直到最后的洗涤完成。这使得 Microcon 过滤器损坏时可以从收集的流出液体中挽救探针。在最后的洗涤完成后，合并装有流出液体的试管。在此阶段加入磷酸盐缓冲液（pH 7.0）至浓度为 10 mmol/L 可以使得三磷酸盐核苷酸降解成单和二磷酸盐形式减至最低。未经纯化的染料可以在－20℃下保存最多一年。不要将染料储存在冰箱（经常开门），因为曝光会加速染料的降解（见本方案结束部分关于最后注意事项的信息栏）。

高盐缓冲液（HSB）

1 mol/L 乙酸三乙胺（pH 8.0）

10％乙腈

将缓冲液通过 0.22 μm 尼龙滤膜过滤。在即将用于 HPLC 之前向缓冲液吹氩气进行脱气。

低盐缓冲液（LSB）

25 mmol/L 乙酸三乙胺（pH 8.0）

10％乙腈

配制 HSB 的 40 倍稀释液，并加入乙腈至 10％。将缓冲液通过 0.22 μm 的尼龙滤膜过滤。在即将用于 HPLC 之前向缓冲液吹氩气进行脱气。

磷酸盐缓冲液（0.5 mol/L 和 10 mmol/L，pH 7.0）

乙酸三乙胺（TEAA）（1 mol/L，pH 8.0）

乙酸三乙胺（TEAA）（15 mmol/L，pH 8.0）

将 0.750 ml HSB 和 49.25 ml H_2O 混合来配制，并将溶液通过 0.22 μm 过滤器。

225

专用设备

HPLC 柱［Dionex NucleoPac 9×250 PA-100；P/N 43011（阴离子交换）］

HPLC 设备（Perkin Elmer 系列 200 LC 泵，Perkin Elmer UV/Vis 检测器 LC 295）

Sep-Pak 柱，C18；每种 Cy 染料需要一根柱子，每三轮顺序操作（20515，Waters Crop.）

注射器（12 ml，每种染料一支）

方法

Sep-Pak 纯化

1. 在50 ml试管中分别准备下列材料：

   乙腈（一组三轮的操作消耗约 40 ml）

   ddH_2O，经 0.22 μm Millipore 过滤器过滤

   15 mmol/L TEAA

2. 上样之前，用 10 ml 乙腈、5 ml H_2O 和 10 ml 15 mmol/L TEAA 依次洗涤Sep-Pak柱。
3. 向样品中加入 1mol/L TEAA，使最终浓度为 15 mmol/L TEAA（3.25 ml 体积中要加入 50 μl）。加入后通常使产量提高 1%～2%。

   柱中最多可上样 150 个 OD_{260} 单位的样品，冻干至体积为 3～5 ml。以 6.5～7.5 ml 的较大体积上样能够得到适当的产量。曾经使用高达 10 ml 的体积，但是产量减少了。

4. 向柱中加样，在接近柱子的顶端可以看见染料。用 5 ml 15 mmol/L TEAA 冲洗柱子。
5. 用 3 ml 乙腈从柱中洗脱 Cy 标记的 dNTP，速度为每秒钟 1 或 2 滴。弃去开始的无色液滴，将有色的液滴收集到 2 ml 的离心管内，并弃去剩余的液滴。

   预期的收集体积为约 1.0～2.0 ml。洗脱后柱子上可能会残留一些颜色。

6. 立即用上样前方案（步骤 2）洗涤柱子。最多可以顺序处理另外两个样品，在样品之间要重新洗涤柱子，并将染料分开收集。

   如果柱子放置 5 min 以上，上样前方案中在用 H_2O 洗涤柱子时可能会有数滴带有颜色的液滴流出，并且产量会减少 5%～10%（或者更多）。柱子重复使用不要超过两次（共使用 3 次），因为产量开始下降。

7. 合并有颜色的样品并测量 UV 吸光度（在每种染料的最大光吸收处）。用染料的消光系数计算产量（见**测量染料标记 DNA 的效率**信息栏）。这个阶段预期的产量为每种 Cy 标记的 dNTP 的 70%～80%。

   一次洗涤——上样——洗涤——洗脱循环通常需要 15 min 来完成，并且需要使用一个环形支架。然而，通过按比例放大柱子的尺寸和使用真空支架装置平行处理多根柱子以增加产量是可能的。这样的柱子和装置可以在 Waters 目录中找到（见本方案结束部分关于**最后注意事项**的信息栏）。

HPLC 纯化

226

8. 在冷冻干燥机中，将从 Sep-Pak 系列纯化（来自步骤 7）得到的合并材料体积减少到约 0.5～1.0 ml，将不同的染料分开。

   这个过程通常需要花费 12～18 h。在冻干过程中，材料有可能完全冻干或者不完全干燥，这是正常的。如果需要，将干燥的团块溶解或者用 0.2 mol/L TEAA 调整溶液的体积使 Cy 标记的 dNTP 的合适浓度为 200 nmol（染料）/μl。在最终溶液中 TEAA 的浓度并不是关键的，在 0.01～0.1 mol/L 范围的任何浓度都是可接受的。

9. 将样品以 14 000 *g* 离心，通过 0.40 μm 的 Microcon 过滤器直到液面到达滤膜。这个过程通常要花费约 10～20 min，取决于样品的浓度。

本步骤清除在样品储存期间由于聚结而形成的高分子物质。

10. 测量滤出液体的光吸收，如果需要，稀释样品至不会超出 HPLC 柱容量（下述的柱子为 10μmol）的浓度。

11. 使用下列缓冲液系统平衡 HPLC 柱，流速为 3 ml/min（适用于一般的清洗方法）

5 min LSB

2 min 梯度过渡到 100% HSB

5 min HSB

2 min 梯度过渡到 60% LSB

5 min 60% LSB

12. 将样品上样到进样环中。

关于样品量和其他问题，见**最后注意事项**的信息栏。

13. 按照下列程序进行梯度洗脱（3 ml/min 的流速，记录仪 0.5 cm/min）。

5 min 60%LSB，40% HSB

10 min 梯度过渡到 80% HSB

1 min 梯度过渡到 100% HSB

10 min 100% HSB

1 min 梯度过渡到 60% LSB

5 min 60% LSB

对应于 Cy 标记 dNTP 的峰应该与其他存在的组分分开至少 1～2 个 1.8 ml 的收集部分。Cy3 标记 dUTP 的保留时间约为 11～12 min，且 Cy5 标记 dUTP 的保留时间为 12.5～13 min。较短的保留时间表明，柱子的有效分离容量低，且柱子接近于使用寿命的终点。在使用的三年时间中，创作本方案的柱子的保留时间缩短超过 20%，同时分离容量也有相当的下降。

最终制备程序：冻干和重新稀释

对大约 150 nmol 的 Cy 标记 dNTP 进行制备操作要能得到约 5 ml 含有 Cy 标记核苷酸的缓冲液（见**最后注意事项**的信息栏）。在本方案的最后步骤中，通过冻干从 Cy 标记的 dNTP 中除出缓冲液中的盐分。

227

14. 将样品从收集管中转移到 50 ml 配有带孔管盖的圆锥形塑料试管中（可以用 18 号针在管盖上刺孔）。

装样量不要超过试管体积的一半，因为在冻干过程中样品会发生溅射并上移。

15. 在干冰上冷冻样品，然后将它们转移到冻干盒中。用铝箔包裹冻干盒以防止样品受到光照。

16. 将样品冻干 6～15 h。

由于有挥发性的乙腈存在，Sep-Pak 纯化后进行的冻干 6～10 h 后可以完成。当样品呈薄膜状和（或）固体残渣颗粒时，冻干已经完成。

17. 将样品重悬在 8～10 ml H_2O 中，然后冻干至干燥。这是一个很长的过程，需

要长达 24 h。

18. 重复步骤 17，然后在 10 mmol/L 磷酸盐缓冲液（pH 7.0）中重悬样品。

    目的在于以 1 nmol/μl 的浓度溶解 Cy 标记的 dNTP（根据对加样到 HPLC 柱上材料数量的估计）。Brown 实验室通常处理数批大约 50～200 nmol Cy 标记的 dNTP，并且在此阶段使用 50 μl 缓冲液重新溶解回收的物质。如果液体太稠，按照每次 500 μl 的量逐步加入缓冲液。回收的 Cy 标记 dNTP 经过定量后（步骤 19），将浓度调节为 1 nmol/μl。

    记住一些 Cy 标记的 dNTP 会黏附在试管壁的上部，并且肉眼可能看不见。为了回收这些残余物质，在试管的上壁滴加液体并转动试管，确保液滴经过了试管壁的全部表面。甚至在可见的残余颗粒重悬之前，液滴常常会带上染料的颜色。

19. 测定样品的光吸收。Cy3 标记 dNTP 的 $A_{560/660}/A_{260}$ 比值应为大约 17，而 Cy5 标记的 dNTP 为 25。较低的比值表明 TEAA 缓冲液仍然存在。

    预期的总体上 Cy3 标记 dNTP 的产量为 50%～70%，且 Cy5 标记 dNTP 为 45%～65%。

最后注意事项

HPLC 柱每批最佳的进样量为含有大约 150 nmol Cy 标记的 dNTP，这依赖于 Microcon 滤出液体的含量和浓度。可以不经 Sep-Pak 纯化直接进行 HPLC 纯化，但是柱子的有效容量将从 150 nmol 减少至少于 50 nmol，而且总体上的产量也会降低。Sep-Pak 纯化后的冻干步骤仍然是必要的，并且可能会花费较长的时间才能完成，因为样品中存在高浓度的缓冲液盐分。

整个方案的限速步骤是 Sep-Pak 纯化步骤，尤其是累积了大量的 Cy 标记的 dNTP 滤出液体时。Waters Corp. 提供较大的 C18 柱和过滤装置，用于同时纯化多根柱子。然而，对洗涤、上样和洗脱体积进行较大的按比例放大时必须要进行优化。

染料产量低的最可能原因有过分暴露在光下、过多的冷冻-融化循环次数以及储存时间超过一年。尽管步骤进行时应尽可能减少直接暴露在光下，本方案可以在实验室灯光环境条件下进行。在长时间的操作过程中（如冻干），或者方案中各步骤之间，试管或容器应该用铝箔包裹或置于暗处（如在密闭的冰桶中）。

将 Cy3 和 Cy5 标记的 dNTP 混合物彼此间分开是可能的，虽然本方案没有针对这个目的进行优化。在对两种修饰了染料的核苷酸分离进行优化时，按照较少数量分开收集洗脱液体，修改起始的 LSB-HSB 混合物，并降低梯度的斜率都是可以考虑的因素。

方案 20

228

杂交和杂交后洗涤

Trey Ideker（怀特黑德生物医学研究院，剑桥，马塞诸塞州，02142）；
Suzanne Ybarra（Eos 生物技术公司，南旧金山，加利福尼亚，94080）；
Sean Grimmond（功能与应用基因组学 ARC-SRC，基因表达谱研究中心，分子生命科学研究所，昆士兰大学圣露西亚校园，昆士兰州，4072，澳大利亚）

与任何杂交反应相同，下面是当标记的靶分子与点样 DNA 微阵列复性时的目标：

• 获得可能的最高信噪比。
• 减少标记的靶分子与具有小部分同源的微阵列单元之间的交叉杂交。
• 确保信号强度正比于标记靶分子的浓度。

这些目标通过优化许多实验参数来实现，包括预杂交和杂交的条件和时间、杂交混合物的体积、离子和缓冲液的条件以及杂交混合物中标记靶分子的含量。关于实验设计上的特殊考虑，请参考第 7 章。

通常，标记靶序列与 DNA 微阵列的杂交条件与常规的 Southern 或 Northern 杂交是类似的。使用高离子强度的缓冲液来减少静电排斥，促进与互补链的结合；使用封闭剂和去污剂来减少背景噪声；在中到高等严格度的条件下进行杂交和清洗来减少交叉杂交。

在杂交过程中，每种固定的靶 DNA 在数量上必须是和它互补的、标记的靶序列的 10 倍或者更高的摩尔数。这种安排保证了在一个较宽的靶序列浓度范围内，杂交过程遵循拟一级反应动力学。任何阵列单元的信号强度取决于在杂交混合物中同源靶序列的起始浓度。因而，靶序列的浓度增加两倍会使得杂交信号的强度也将增加两倍，并且高丰度的 RNA 会产生比低丰度的 RNA 更强的信号。当固定的靶 DNA 的数量有限时，系统的动态范围受到限制，并且对差异表达的估计将缩减（Holler et al. 1997；Yue et al. 2001）。在表达分析中，通常将靶序列（cDNA 或 PCR 产物）以 100～500 ng/pl 的浓度点样到微阵列上（Schena et al. 1996；Welford et al. 1998；Deyholos and Galbraith 2001）。

当信噪比高于 1.5～2.0 时，微阵列特征的点阵图样变得可以检测。微阵列实验灵敏度一种有用的量度是，能够通过与和它同源的阵列上的 DNA 单元杂交后被可靠地检测到的最少部分的 mRNA（Bertucci et al. 1999）。在一个理想的实验中，这个最少部分检测限为 1/1 000 000～2/1 000 000——足够用于检测每个细胞中仅存在数个分子的 RNA。

229

如果杂交混合物中靶序列的浓度过低，复性将会很慢并且减弱的信号有可能不会被荧光扫描仪检测到。杂交时需要的最低数量的靶序列应该是在样品中刚好含有足够数量的最低丰度 RNA 分子，在杂交以后能够产生可读出的信号。由于随着 RNA 数量的增加，信噪比和动态范围都随之升高，通常在杂交混合物中含有的靶序列是计算的最低数量的3～4 倍。然而，如果靶序列的浓度过高，如此计算的最低数量大 10 倍，信号的动态范围会下降，并且信噪比可能会变坏或者没有改善（Yue et al. 2001）。作为一个经验规则，在标准的（15～35 μl）微阵列杂交混合物中，使用的标记靶分子要等价于至少 1 μg poly(A)$^+$ RNA（大约 100 μg 总 RNA）。如果不能得到这么多量，那么 RNA 制备物在标记前应该进行扩增（Guatelli et al. 1990；Van Gelder et al. 1990；Eberwine et al. 1992；Luo et al. 1999）。Poly(A)$^+$ RNA(1 μg) 通常对应于 10～40 pmol 未掺入的荧光染料。

本方案描述了微阵列的预杂交、杂交和清洗步骤。在主要方案的最后有一个关于杂交的附加方案，使用微阵列载玻片的叠放。

开始之前：准备清洁的作业面

灰尘是微阵列实验中是最常见的污染之一。在开始杂交反应之前，需清洁并用无尘垫子覆盖全部实验台的工作区域。有两种常用的无尘垫子，包括铝箔和吸尘的工作台垫子，塑料一面向上。按照一般方式使用工作台垫子（吸尘的一面向上）会带来纤维和尘埃的污染源，通过静电作用附着在洁净干燥的载玻片上。戴上无粉的丁腈橡胶手套。避免标记的靶序列暴露在强光之下。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。
使用前将储存液稀释到合适的浓度。

牛血清白蛋白（BSA）（10%，*m/V*，Worthington）

50×Denhardt 溶液

甲酰胺，分子生物学级（100%）（BP227-100，Fisher Scientific）〈!〉

见关于**甲酰胺**的信息栏。

异丙醇（HPLC 级）

SDS(20%)（US75932，USB）〈!〉

20×SSC(pH 7.0)（US 19629，USB）或 20×SSPE

核酸和寡核苷酸

230

C_0t-1 DNA(20 mg/ml，人或小鼠的最佳)（GIBCO）

Poly(A)（10 mg/ml）（P9403，Sigma）

靶 DNA，按照方案 13 所述方法标记〈!〉

tRNA(10 mg/ml)（15401-011，GIBCO）

专用设备

盖玻片（如 HS40、HybriSlips、Grace BioLabs，或 Corning Cover Glass，2935-246，Fisher Scientific，12-548-5P）

重要事项：清洁盖玻片

- 将盖玻片短暂地浸入 dH_2O，然后在另一个装有 dH_2O 的容器中蘸洗数次，最后在装有 100% 异丙醇的容器中蘸洗数次。
- 用 Kimwipe 纸巾擦去异丙醇。
- 用压缩空气吹净灰尘。检查盖玻片相对洁净且没有灰尘。

Coplin 染缸（S 5516，Sigma）或者显微镜载玻片染色盘（VWR 25461-003）

避光清洗容器

清洗容器可以是 Coplin 染缸（见上）、覆盖有铝箔带的大培养皿，或者是带有匹配载玻片架（VWR 13197-272）的容器或染色皿（VWR 13197-274）

微阵列扫描仪（表 3-9）

微量浓缩器（如 Microcon YM-100）

Microduster Ⅲ(8 oz. 适合于再次填充的；VWR 21899-092)

载玻片架（900234，Wheaton）

方法

预杂交

预杂交用于：

- 在加入靶序列之前从载玻片上洗去未结合的 DNA。
- 封闭载玻片表面可能与标记的靶 DNA 非特异性结合的反应性基团（如自由氨基）。

预杂交可以降低非特异的背景，但是增加了在操作过程中载玻片损坏或者污染的风险。尽管关于该处理对多聚赖氨酸玻片的作用存在各种观点，在 γ-氨丙基硅烷（GAPS）玻片（Corning Microarray Technology）上推荐使用预杂交，但是 Telechem SuperAmine 载玻片并非一定需要预杂交；一些研究人员发现这些玻片预杂交是有益的。

1. 在室温下，将玻片放入装有 0.1%～1.0%(*m*/*V*) BSA，5×SSC，0.1% SDS 溶液的 Coplin 染缸中。将 Coplin 染缸在 42℃水浴中放置 30～45 min。

   所有的预杂交和杂交后洗涤都在 Coplin 染缸（如这里所述）或者显微镜载玻片染色盘中进行。另一种可供选择的玻片准备方法在步骤 4 下面的信息栏中给出。

2. 在室温下清洗载玻片，换水数次。
3. 用 100%异丙醇（HPLC 级）漂洗载玻片。

   一些研究人员使用异丙醇清洗玻片对杂交起相反作用。如果有这个问题，请使用下面的替代方案。

4. 在加入杂交溶液之前，在空气中或者通过离心（室温下以 2000 *g* 离心 5 min，有阵列的一面朝外）干燥玻片。

   将玻片立即用于杂交。如果玻片干燥时间超过 1 h，杂交效率可能会迅速下降。

下面的各步替代方案可以用于代替步骤 1～4 进行载玻片的清洁：

- 在 55℃下，将玻片浸入 5×SSC/0.2% SDS(pH 7.0）中约 40 min。温育时间很大程度上取决于使用的载玻片的批次。使用一个 Coplin 染缸，放在 55℃的水浴锅中，无需搅动。
- 在室温下迅速将玻片在 dH_2O 中蘸 5 次。
- 用氮气吹干或者离心干燥玻片。

标记靶序列的杂交

杂交缓冲液通常含有：

- 2.5～5×SSC 或 SSPE。SSPE 具有更好的缓冲容量，但 SSC 是更常用的。在含有常呈酸性的甲酰胺的缓冲液中必须使用 SSPE。
- 0.1%～0.25%SDS

此外，杂交缓冲液通常含有以下成分的组合。在配制杂交溶液之前，用 0.22 μm 过滤器过滤所有的组分。

- 50%甲酰胺。加入甲酰胺可以降低杂交体的熔点，并使杂交可以在更低的温度下进行，从而减少了蒸发的倾向。用含甲酰胺的溶液杂交可在 42℃下进行，相比之下，用水溶液杂交的温度为 65℃。
- tRNA(0.75 mg/ml)*
- C_0t-1 DNA(1.0 mg/ml)*
- poly(A)（1 mg/ml)*
- BSA(0.2 mg/ml）或 5×Denhardt 溶液
- 鲑鱼精 DNA(100 μg/ml)

*这些封闭试剂通常组合使用。一些研究人员选择刚好在最后的洗涤（方案 13，步骤 8）之前向标记的靶 DNA 中加入这些封闭试剂以减少由于超滤膜的吸附而造成的损失。一个典型的杂交混合物可能含有 50%甲酰胺、5×SSC、0.1%SDS、5×Denhardt溶液。

5. 将适量 Cy3 和 Cy5 标记的靶 DNA 混合在一起，并按照方案 13 中步骤 8 和步骤 9 所述方法用 Microcon YM-100 装置将混合物浓缩至最终体积大约 6 μl。
6. 配制含有靶 DNA 和下列成分的杂交溶液，按照如下体积（24 mm×24 mm 盖玻片要约 16 μl，24 mm×40 mm 盖玻片要约 24 μl，24 mm×60 mm 盖玻片要约 32 μl)：

232

20×SSPE	10 μl
甲酰胺	20 μl
C_0t-1 DNA	2.0 μl
poly（A)	2.0 μl
SDS（10%)	0.5 μl
标记靶序列	约 6.0 μl

以上为典型杂交溶液的配方；然而，如同步骤 5 前面的信息栏中所述，组分可能会变化。

7. （可选的）将最终的混合物在微量离心机中以 10 000 *g* 离心 5 min，然后将上清液转移到新的试管中。

   一些研究人员增加离心步骤以除去靶序列纯化过程中带入的微量玻璃纤维和其他高分子质量的颗粒。

8. 将上清液在 100℃下加热 2 min。将溶液在 30℃水浴中冷却 30 s。

   加热杂交混合物可以降低背景。不要将溶液放置在冰上，因为 SDS 会发生沉淀。

9. 将适量的杂交溶液加到 DNA 微阵列上。在载玻片上小心地盖上盖玻片。

   建立杂交反应是微阵列方法中具有较大技术挑战性的步骤之一。这是一个复杂的过程，需要在空白的玻片上练习。杂交溶液必须均匀的加到微阵列的表面，并且一旦加好杂交溶液，就必须防止蒸发，通常的做法是在溶液上盖上盖玻片，然后在一个潮湿的盒中温育微阵列。由于这些步骤发生在微阵列方案接近结束的部分，紧接在微阵列制备和cDNA标记之后，任何相关的问题都可能导致数小时的前期实验工作被浪费掉，或者至少会对微阵列图像结果的质量产生重大影响。受这些因素影响，最近市场上出现了许多自动处理“工作站”，试图将杂交和清洗步骤自动化。然而，这些设备仍需改进和发展，

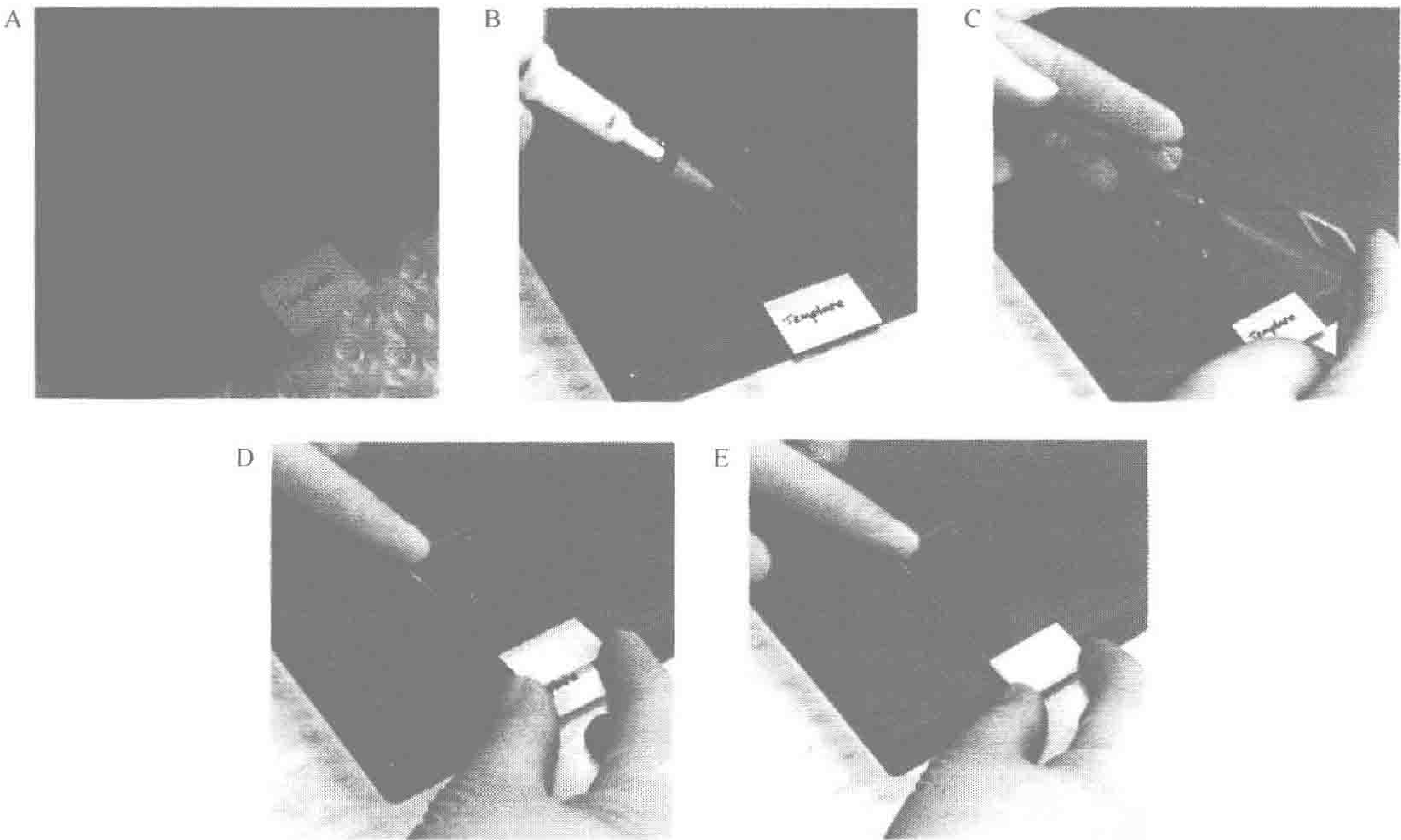
我们目前得到的最佳质量的微阵列图像都是通过仔细地手工加入杂交液和放置盖玻片而获得的。图 3-22 和图 3-23 中描述了两种方法。两者都可以达到同样的结果，即一旦被掌握，它们会产生重复性相当好的结果。选择何种方法取决于个人的偏好。

不要向微阵列中央加入大量的靶序列，这会导致在盖玻片下方产生气泡。

如果杂交溶液的量太多，盖玻片会浮在载玻片的表面。下面继续处理载玻片时（如将它放入杂交盒或者带着它走到水浴），盖玻片不能够继续盖在阵列点阵区域上方。杂交溶液位于点样区域中心时，直接放下盖玻片是很难操作的。与此相反，杂交溶液应该加到靠近点样区域内缘的左侧，这样当盖玻片缓慢放下时杂交溶液便可以取代空气。

233

标准的组织学盖玻片对于微阵列实验而言是不够洁净的，必须按照材料部分中的信息栏所述的方法进行清洁。不干净的盖玻片在建立杂交时会产生气泡，导致高的背景和（或）黑洞（见本方案结束部分的疑难问题解答指南）。微小的气泡通常不会造成问题，在溶液加温时会消失。试图通过移动盖玻片来赶走或者移动小的气泡的操作会带来危害。



234

图 3-22　建立杂交反应：手工将杂交溶液从盖玻片转移到倒置的微阵列中

将一张合适的盖玻片放在模板载玻片上（A）并向盖玻片的中央加上杂交溶液（B），仔细操作避免产生气泡。在盖玻片上方，将带有点样和处理过的阵列的玻片倒置（C），轻轻放下载玻片，直至它和杂交混合物接触（D）并且由盖玻片转移到微阵列载玻片上（E）。倒置微阵列玻片，使得盖玻片位于上表面。小心不要让微阵列玻片和模板玻片完全接触，否则靶分子溶液会使两张玻片粘在一起，无法分开

10. 由于杂交是在很小的体积下进行的，在密闭的潮湿环境下温育以防脱水干燥是非常重要的。使用以下任何一种安排都会取得满意的结果：
   - 一个可以放 25 张载玻片的塑料玻片盒（不透光的），其中放置一张潮湿的 3M Whatman 滤纸（25 张玻片盒使用 1 ml 3×SSC）。
   - 在一个空的移液器吸头盒中，底部装有 5 ml 3×SSC，将载玻片平放在有孔的表面。

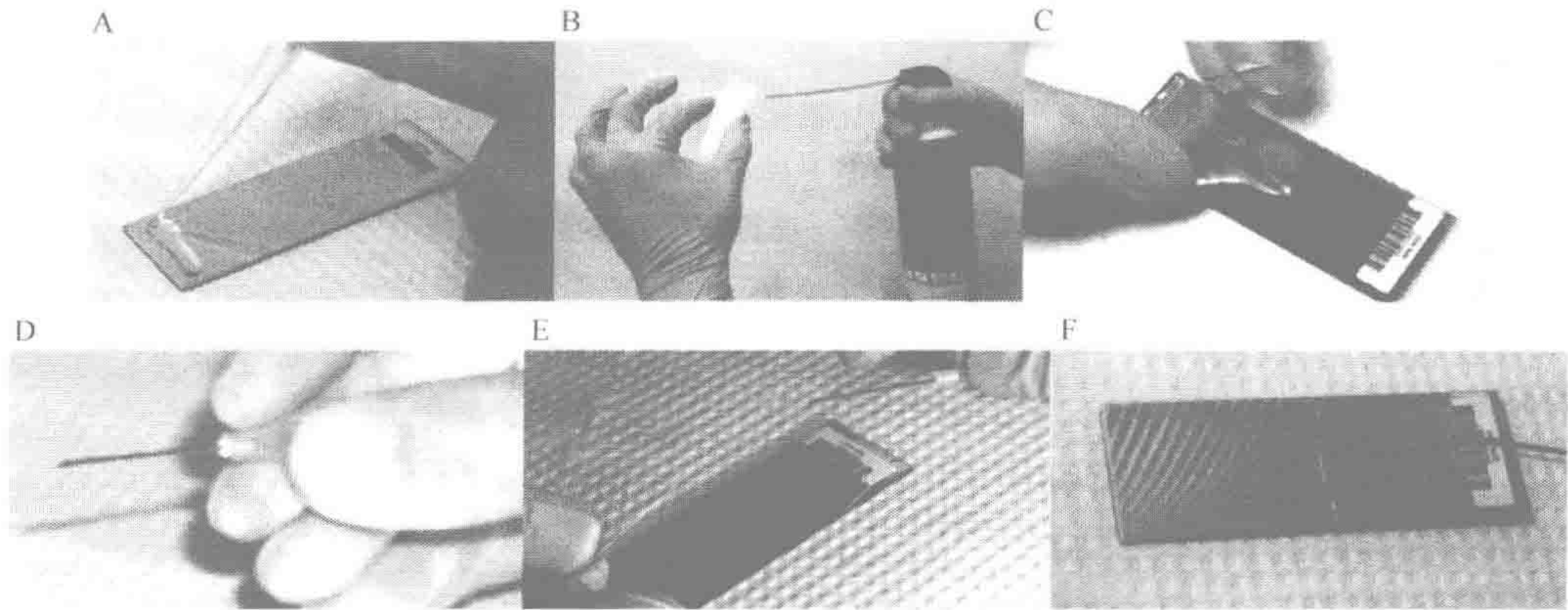


图 3-23　建立杂交反应：杂交之前手工加入杂交溶液和放置盖玻片

图版 A～F 中的图片示例了使用印刷在覆盖了反光材料的标准 25 mm×70 mm(1×3 英寸) 载玻片 (APBiotech Type 7 RPK0328) 上的微阵列建立杂交反应。开始时，用压缩氮气吹去微阵列载玻片表面的灰尘，然后将载玻片平放在洁净的工作台上。使用装有标准的 P200 吸头的微量移液器吸取约 40～50 μl 的杂交溶液（标记 cDNA 样品溶解在杂交缓冲液中）并以均匀的液滴形式加到微阵列载玻片靠近一端的边缘上（A）。需要的确切体积取决于载玻片特定的表面化学性质和所使用的杂交缓冲液，目标是使溶液均匀的扩展在整个微阵列的表面而不会在边缘浪费溶液。在这一过程中，偶尔会从吸头挤出一个或者数个气泡到杂交溶液的表面。由于这些气泡后来会驻留在盖玻片的下表面并干扰杂交，因而要在继续实验之前除去这些气泡。使用小到中号针的针头很容易破坏气泡（D）。也可以将气泡直接吸回吸头，但是这种方法不可避免地将一些杂交溶液本身吸掉（并浪费）。通过练习，能够用这种方法移取溶液，而且完全不会带来气泡。加完杂交溶液的液滴后，立即取一片洁净的盖玻片并用压缩氮气短时吹净表面灰尘数次（B）。盖玻片应该足够大，要能够完全覆盖载玻片上含有微阵列的区域，从边缘向中心拿住盖玻片以减少氮气的吹力使盖玻片破碎的可能性。将它与微阵列载玻片以相同的方向对齐，但是翘起一些角度并且不要接触到杂交溶液的液滴（C）。从盖玻片位于载玻片后方的最低边缘开始，缓慢拖动盖玻片向前移动，使边缘和液滴接触。一旦开始接触，让盖玻片的最低边缘停在载玻片上。现在，缓慢地换用小到中号针（D 中显示的是 Becton-Dickinson 20G11/2 针）的针头托住盖玻片抬起一端的边沿。如果针尖是斜面的，针头较长的一侧应该朝下。为了确保更换过程平缓地进行，将针的底部（图中的黄色塑料）牢固地靠在工作台上。以此为支点转动针，使针尖向上支撑住盖玻片抬高一侧的边沿（E）。手指和拇指松开放下盖玻片，使得不需要手扶，它的一端由微阵列载玻片支持，并且另一端用针尖支持。针的基部固定在工作台上，慢慢降低针尖，使盖玻片的抬高端跟着下降。将盖玻片的内表面靠近微阵列载玻片时，杂交溶液开始在两者之间由于毛细作用水平展开。小心不要让盖玻片掉下来，而是要继续降低盖玻片，直至它几乎完全平行并停留在载玻片表面。此时，杂交溶液已经通过载玻片的 1/2 到 2/3（F）。当继续下降针头变得非常困难时，从盖玻片的下方抽出针头。不要担心针头会划伤载玻片的表面，它对微阵列几乎没有损害。杂交溶液将迅速前进并填满盖玻片与载玻片之间残留的空隙。如果溶液没有扩展得足够远，或者从盖玻片的边缘溢出，尝试在过程开始时增加或者减少加到载玻片上杂交溶液的量。注意：在载玻片的一端加入溶液并让它沿着载玻片的长轴方向扩展（沿着载玻片的长轴方向加入溶液并让它沿着短轴方向扩展进行比较）可以缓慢而且更为精确地放下盖玻片，并能够减少扩展过程中气泡产生的机会

- Telechem 或 Corning 的商品化杂交盒，载玻片周围完全密封，因此可以浸入水浴中。

  制作适用于单张载玻片杂交的杂交盒的详细方法可以公开获取，见 http://cmgm.stanford.edu/pbrown/mguide/index.html.

11. 在 42℃（如果杂交溶液中含有甲酰胺）或者 65℃（使用水配制的杂交溶液）下将载玻片温育 14～16 h。

    保持合适的杂交湿度非常重要。过分干燥，盖玻片会黏附到微阵列上，可见到高的背景；过分潮湿，冷凝的液体会溢进盖玻片周围的区域，降低信号的强度。

芯片杂交后的清洗

重要事项：在清洗过程的任何阶段都不要让载玻片变得干燥。不要使用过多的 SDS，在扫描之前要彻底除去 SDS，以避免高的背景。通常在室温下进行清洗，尽管一些研究人员在最高 55℃下进行洗涤步骤 1 和步骤 2。一组特定靶序列的最佳清洗温度要根据经验来确定。

载玻片清洗的同时，打开扫描仪进行预热。

12. 配制下列洗涤液，并将它们用 0.22 μm 一次性过滤器进行过滤。

    清洗 1：0.5×SSC，0.01%SDS

    清洗 2：0.5×SSC

    清洗 3：0.1×SSC

    清洗 4：0.01×SSC

13. 拆卸杂交盒。如果杂交盒是浸入水浴中的，在松开螺丝之前，仔细擦干水痕，尤其是两个半片盒子之间的水痕。
14. 向一个 Coplin 染缸中加入清洗 1。从杂交盒中取出载玻片，然后立即浸入到清洗 1 中，直到盖玻片分离。

    除去盖玻片时可能需要轻轻搅动。不要试图手工移去盖玻片。如果有一些阻碍并且盖玻片看上去似乎不动，将载玻片放回玻片盒中，在载玻片上方倒一些清洗 1。放置大约 1 min，然后再尝试。每张载玻片使用新的清洗 1。

236

15. 一旦盖玻片除去后，将载玻片放到玻片夹持器上，然后浸入装有约 250 ml 清洗 1 的避光容器中。将容器放在轨道振荡器上，并将载玻片振荡 2 min。
16. 将载玻片和夹持器一起浸入到另一个装有约 250 ml 清洗 2 的容器中，并且再次将载玻片在振荡器上振动 3 min。
17. 将载玻片和夹持器一起浸入到另一个装有约 250 ml 清洗 3 的容器中，并且再次将载玻片在振荡器上振动 3 min。
18. 重复步骤 17 两次，每次清洗 1 min。
19. 快速（<10 s）将载玻片和夹持器一起浸入到另一个装有约 250 ml 清洗 4 的容器中，立即将载玻片放到铺有两层 3M Whatman 滤纸的离心机微量滴定板夹持器中。低速离心约 5 min 迅速干燥玻片。
20. 将载玻片放置在避光的盒子内直到准备扫描。将微阵列在当日进行扫描，因为随着时间的延长信号会降低。

    关于微阵列杂交后扫描/成像的详细信息，见表 3-9 和下面关于微阵列阅读器的信息栏。

微阵列阅读器

两种类型的仪器可用于靶序列和微阵列杂交后数据的采集。共聚焦显微镜，又称作激光扫描仪，依赖于使用一个或多个激光器作为激发光源，采用逐点的方式对微阵列进行扫描。在该系统中，光电倍增管用作检测器。激光可以依次或者同时扫描阵列样品。相反，阅读阵列的成像系统利用电荷耦联装置（CCD）照相机。基于 CCD 的成像仪器通常利用一束白光作为激发光源，通过滤色块或者滤色转轮，并从阵列以逐区的方式采集一系列的图像。在表 3-9 中介绍了各种微阵列阅读器的供应商和特点。关于这些系统的最新信息，请参考供应商的网站。

表 3-9 各种阵列扫描仪的特性

公司	基本设计	分辨率/μm	10K 玻片的扫描速度	灵敏度	激光和激发范围	相关软件	其他特点
Affymetrix (www.affymetrix.com) 418 Array Scanner	旋转激光扫描显微镜	10	扫描带有 10 000 个样点的玻片小于2 min	小于 1 个荧光基团/μm^2(信噪比大于 3.5)	红色二极管，激发波长 635 nm，发射滤色镜 665 nm；绿色倍频 YAG 激发波长 532 nm，发射滤色镜 570 nm；第三个发射滤色镜可以检测最多 10 种染料	Jaguar 2.0 软件集成了仪器控制和数据分析；MicroDB 2.0 和 DMT 2.0 可进行数据挖掘	线性动态范围 3～5 个数量级；集成了两个附加的激光器和三个附加滤色镜
Agilent (www.chem.agilent.com)	扫描激光	5	每张玻片小于 8 min	小于 1 个荧光基团/μm^2	SHG-YAG 激光器，532 nm HeNe 激光器，633 nm	PC 数据系统、数据提取和数据分析软件	扫描仪的焦距可以连续调节
Amersham (molecular Dynamics) (www.apbiotech.com) Typhoon 9400	可以对磷光、荧光、化学发光扫描和(或)成像来检测	10			红光激发荧光在 633 nm；绿光激发荧光在 532 nm；蓝光激发荧光在 457 nm 和 488 nm	ImageQuant 图像分析软件	线性动态范围超过 5 个数量级
Alpha Innotech (www.alphainnotech.com) AlphaArray 扫描仪	基于 CCD 的成像系统	整个阵列	小于 1.5 s	0.06 个荧光基团/μm^2	8 个激发波长来检测多种荧光基团	图像和数据分析软件	高通量成像仪 光学管理系统(NovaRay 光学管理)减少采集时间
Axon Instruments (www.axon.com) GenePix 4000B	扫描激光	5～100	5 μm 时需要 12 min	Cy3 和 Cy5 为 0.1 个荧光基团/μm^2	409 nm，473 nm，532 nm，594 nm，635 nm，670 nm，690 nm	GenePix Pro 3.0 阵列分析软件	可提供小型设备 (13.5 in×8.75 in×17.5 in)；焦距可调

续表

公　司	基本设计	分辨率/μm	10K 玻片的扫描速度	灵敏度	激光和激发范围	相关软件	其他特点
GeneFocus (www. genefocuscom) DNAscope IV, V	扫描共焦激光束	5、10、20	10 μm 时需要 6 min	小于 1 个荧光基团/μm^2	532 nm, 635 nm, 488 nm	MACROview	最多 3 个激光器，四组滤色镜
DNAscope LM	扫描共焦激光束	5、10、20	10 μm 时需要 6 min	小于 1 个荧光基团/μm^2	532 nm, 635 nm	MACROview	最多 3 个激光器，四组滤色镜
Packard BioScience (www. packardbiochip. com) ScanArray Lite ScanArray 4000 ScanArray 5000	激光共焦光学系统	5、10、20、30、50	整个扫描区域小于或等于 5 min	小于 0.1 个荧光基团/μm^2	可选激光器：488 nm, 514 nm, 543 nm, 594 nm, 612 nm 和 633 nm	QuanArray 微阵列分析软件	多达 11 种发射滤色镜可供选择；条形码和容量为 20 张玻片的自动装载器可选
Cyclone Gene Array System	磷光成像仪，带有螺旋扫描设计	42	3～10 min	小于 2 dpm/mm^2/h 动态范围 5 个数量级		OptiQuant 和 QuantArray 微阵列分析软件	成像面积：240～537.5 cm^2 放射性同位素标记
Perkin Elmer (www. perkinelmer. com) GenTac LSIV	非共焦式扫描激光	2～30（用户自定义）	1 s～10 min（预设定或自动）	0.1 个荧光基团/μm^2	4 种激发激光器和 2 个附加的光电倍增管	Microsoft SQL server 7.0-db 引擎	高通量，一次可以同时对最多 24 个微阵列成像；暗视场照明，荧光滤色镜可提供高质量图像，背景低
Virtek Vision Corp (www. virtekbiotech. com) Chipreader	扫描激光共焦；2 个 PMT 检测器；激光器和 PMTs 独立控制	10～80 (Chipreader) 5～40（高分辨） 3～24（极限值）	25 条扫描线/s；分辨率 10 μm 时完全扫描用 6 min	16 位动态范围	532 nm 和 635 nm	CR 2.0	可提供小型设备（11 in. ×9 in. ×12 in.）

附加方案：微阵列玻片的叠放

237

Sean Grimmond

ARC-SRC for Functional and Applied Genomics, Gene Expression Profiling Facility, Institute of Molecular Bioscience, University of Queensland St Lucia, QLD, 4072, Australia

玻片的叠放用于将由垫片隔开的两块微阵列玻片同时与一份靶序列进行杂交。两块玻片面对面地叠放在一起，位于上方的玻片偏移约 2 mm 以形成一个边缘，便于向其上面加入杂交溶液。这种玻片排列方式形成一个空腔，可以向其中注入杂交溶液。很多商品化的微阵列基底在一端预先标记了黏性条形码，它们是阵列叠放理想的垫片。

起初，对一端带有不干胶条形码标签的 CMT GAPS 玻片进行叠放，这些条形码成为玻片之间的垫片。此后，CMT GAPS 玻片的条形码发生改变而嵌入玻片内部，这些玻片仍然可以用于玻片的叠放。要注意许多商品化的杂交盒不能容纳叠放的微阵列玻片。

方法

1. 根据条形码（或其他垫片）对齐两张玻片，使它们从头到尾对齐排列，点样有微阵列的一面朝上。将一张玻片翻转到另一张上面，使得点样有微阵列的一面相对。将上方的玻片滑动 1～2 mm 以便于加入杂交溶液。
2. 按照主要实验方案的步骤 6 配制标记靶序列和杂交溶液。叠放式杂交反应需要大量的（通常大约 80～100 μl）杂交溶液。

   如果使用了带有黏性不干胶条形码的玻片，需要的杂交体积为 80～100 μl。如果玻片不带有条形码（或有嵌入式的条形码），则需要 40～50 μl 的杂交体积。
3. 在两张玻片间隔处离玻片末端约 1/3 的地方，通过单次移液操作加入杂交溶液。一次性加入杂交溶液防止了气泡注入。当全部杂交溶液都加入完毕后，用镊子轻按两张玻片使它们互相平齐以进行杂交反应。避免接触叠放玻片的边缘，因为毛细作用会将杂交溶液从其中吸走。

疑难解答指南*

PART Ⅰ：点样阵列疑难解答（见彩图）

240

图像 1：理想图像
图像 2：气泡
图像 3：慧尾
图像 4：损伤的基底
图像 5：扩大的样点
图像 6：圆环样
图像 7：边缘干燥
图像 8：边缘褪色

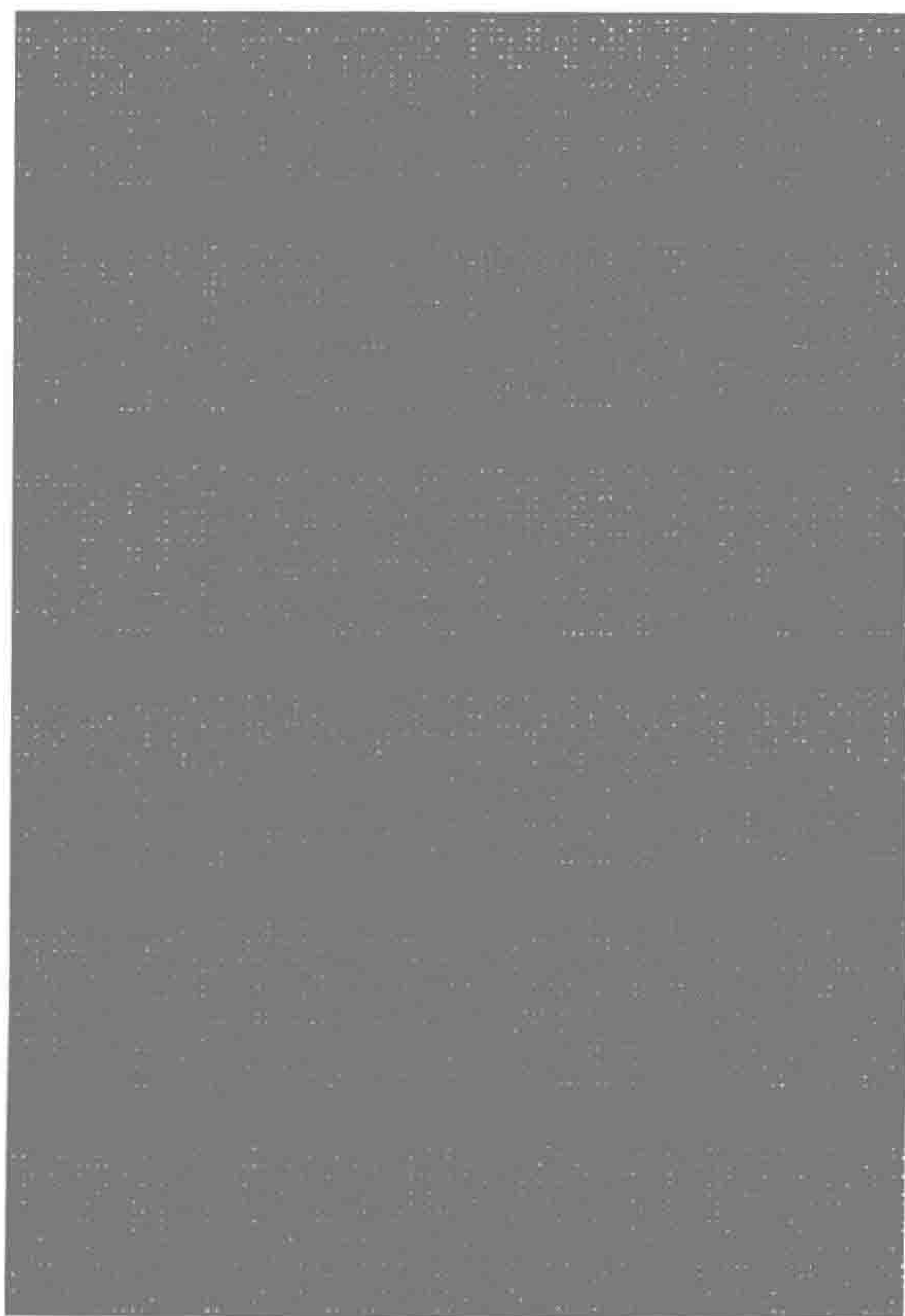
* 由 carolyn Best，Trey Ideker，Javed Khan，Damares Monte，Ed Tom 提供。

图像 9：高背景：荧光
图像 10：高背景：黑洞
图像 11：不规则样点形态
图像 12：信号强度低
图像 13：颗粒污染
图像 14：点样针堵塞
图像 15：扫描仪故障
图像 16：日复一日点样中的差异

PART Ⅱ：AFFYMETRIX 阵列疑难解答

图像 17：高度不规则的背景
图像 18：颗粒污染
图像 19：明亮的斑块
图像 20：非特异性信号
图像 21：扫描仪故障：信号缺失
图像 22：芯片缺陷
图像 23：样点表面刮痕

241



图像 1

图像 1：理想图像

图像动态范围超过两个数量级（样点强度范围从 200～50 000 个计数）；最小的背景荧光（Cy3 和 Cy5 通道均小于 200 强度计数）；动态范围和背景强度有很强的空间一致性（即从图像的一侧到另一侧）；圆形的，均匀分布的样点，以及几乎没有灰尘、擦痕、污点或其他人为假象。

图像 2：气泡

问题

杂交过程中盖玻片下方大的气泡排斥了靶序列，导致该区域内的信号微弱。气泡部分地或者完全地阻止了与它们附近样点的杂交。气泡也会导致大量荧光在其周边积累。注意：在建立杂交时产生的小气泡随着杂交过程中玻片的升温通常会消失。

原因

- 建立杂交的技术有缺陷。
- 杂交溶液中存在微粒物质。
- 盖玻片不洁净。

可能的解决方法

- 练习建立杂交的技术（图 3-22，图 3-23）。
- 使用前用水和乙醇清洁盖玻片。
- 使用前用不含油的压缩 N_2 清洁玻片和盖玻片。
- 使用微阵列专用的盖玻片，如 Hybrislips。
- 加到微阵列上之前将杂交溶液离心。



图像 2

242

图像 3：彗尾

243

问题

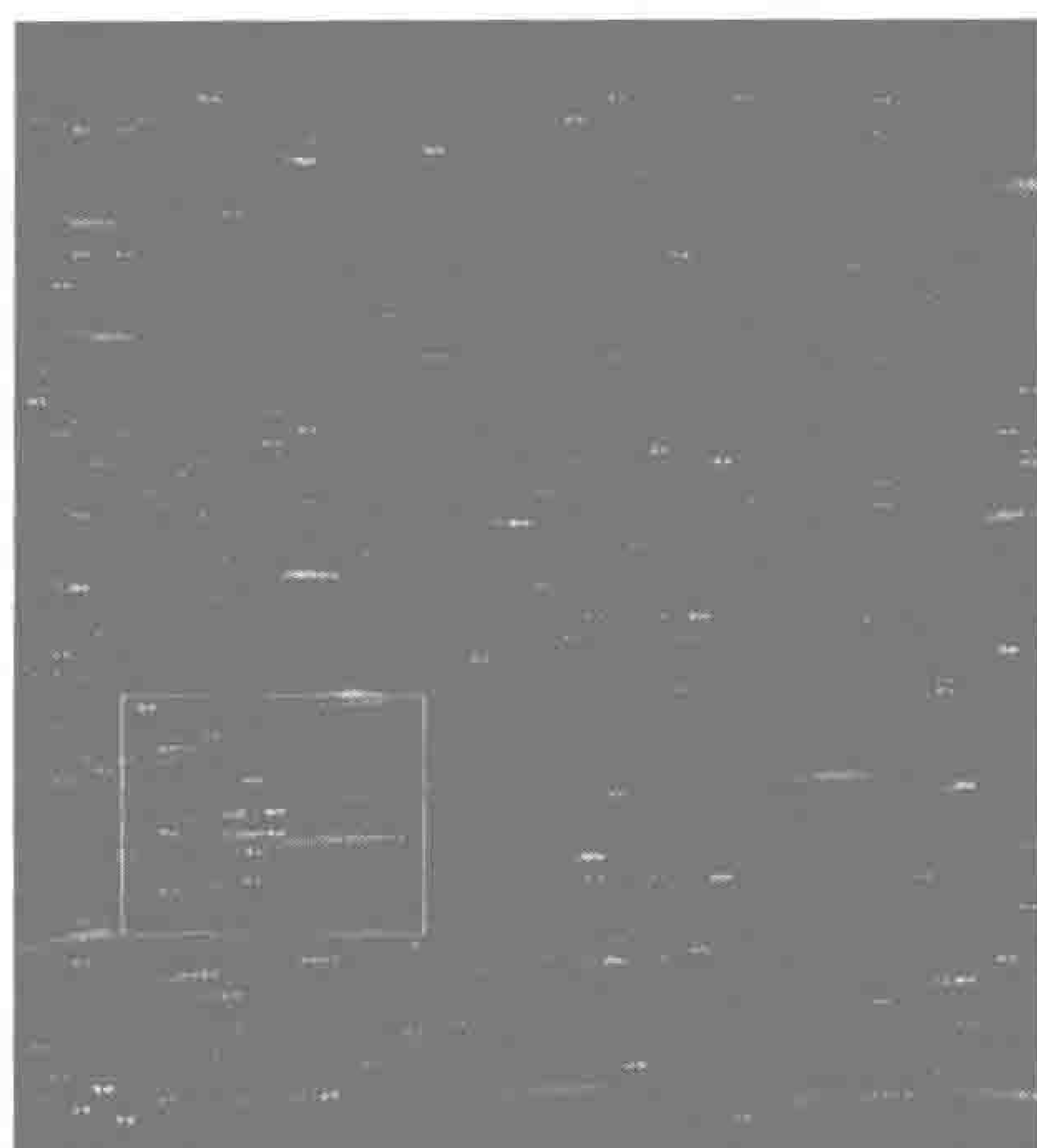
彗尾是 DNA 在样点处发生扩散。DNA 在整块玻片上自右向左拖尾扩散（上图）。彗尾沿着相同的方向扩散，DNA 可能是在封闭过程中或者点样后清洗时发生解离。对该玻片的问题追溯到有缺陷的紫外交联。中图所示方框区域的详细情况。在某些位置，样点呈现黑洞样轮廓，这是因为解离后的 DNA 与这些区域的结合被先前沉积的 DNA 所阻遏。下图大体上具有良好的样点形态，但是其中一个样品产生严重的彗尾，这是由于 DNA 浓度过高造成的。要注意其余的样点中有一部分存在轻微的扩散，而且欠佳的固定使得这个问题变得严重。

原因

- 由于阵列材料中 DNA 的浓度过高导致超出玻片的结合能力。
- 紫外交联不够或者烘焙不充分导致 DNA 在玻片上的固定不充分。
- 点样后清洗技术有缺陷。
- 该批次玻片有缺陷。
- 明亮的样点更容易显现出彗尾。这一问题不总是要归结为过量的 DNA。更为准确地说，它可能是玻片上所有样点的共性问题，而在那些高度表达并且产生强信号从而揭示被取代 DNA 的基因中最容易显现。

可能的解决方法

- 将用于阵列点样的 RCR 产物以 150～250 $\mu g/ml$ 的





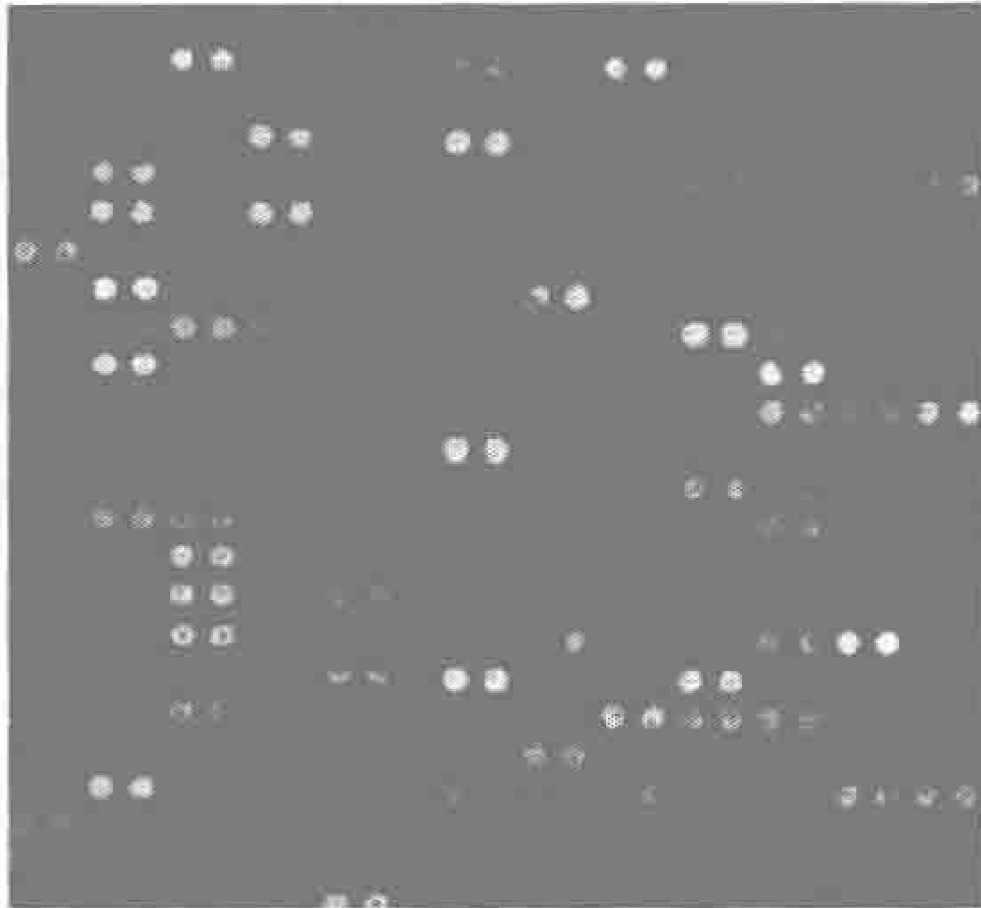


图像 3

浓度重悬。

- 一个普遍性的问题（上图和中图）更可能是由于固定的不充分。检查烘箱的温度。更换和校准交联仪中的紫外灯管。
- 如果玻片上只有少数样点（下图）受到影响，那么这个问题更可能是由于那些样品中过量的DNA造成的。核对DNA的水平（见第1章，方案4）。
- 在点样后清洗时，将玻片迅速浸入清洗溶液内并进行振荡。
- 检查该批次玻片的质量。使用其他替代的玻片基底。

244 图像4：损伤的基底



图像4

问题

对基底造成损伤。（上图）盖玻片重定位时在载玻片基底上产生划痕。（下图）点样针使聚赖氨酸基底产生碎裂而导致的红色杂交背景（箭头）。

原因

- 使用中移动盖玻片，或者在清洗中受到外力时造成的划痕。
- 由于针尖本身的设计或者针尖与玻片表面过快接触而导致的基底碎裂。

可能的解决方法

划痕

- 不要对盖玻片进行重新定位，但是如果有必要的话，从边缘而不是顶部轻轻用力使盖玻片滑动。
- 避免载玻片过分干燥，这会使清洗中除去盖玻片变得困难。

碎裂

- 如果碎裂主要是和一根点样针有关，检查该针尖是否损坏。
- 通常，要减小接触速度和（或）印刷距离；
- 考虑使用更加耐磨损的载玻片基底。聚赖氨酸相当脆弱。

245 图像5：扩大的样点

问题

在点样过程中递送了过多的DNA，导致出现大的不规则样点。在一轮点样过程开始时的那些玻片上情况可能更糟。图中显示了一个大体上较好的微阵列，但是偶尔有一些具有大的不规则样点形态的探针（方框中区域），它们遮掩了相邻的样点。

原因

偶尔出现的样点

- 加样板上孔内液体的量过多。
- 点样针受到轻度的污染或者损伤。

与单个大方格有关的问题

- 点样针的外部受到污染或者损坏并带有过多的样品。

常见问题

- 加样板的大部分或全部板孔装样过满溢出。
- 在点样过程开始时蘸墨不充分（鹅毛笔针）。

- 从加样板中加样时或者与载玻片接触时过长的停留时间。
- 点样时过度的垂直加速。
- 不相容的点样针/表面化学，如疏水系数的不一致。
- 过分潮湿。

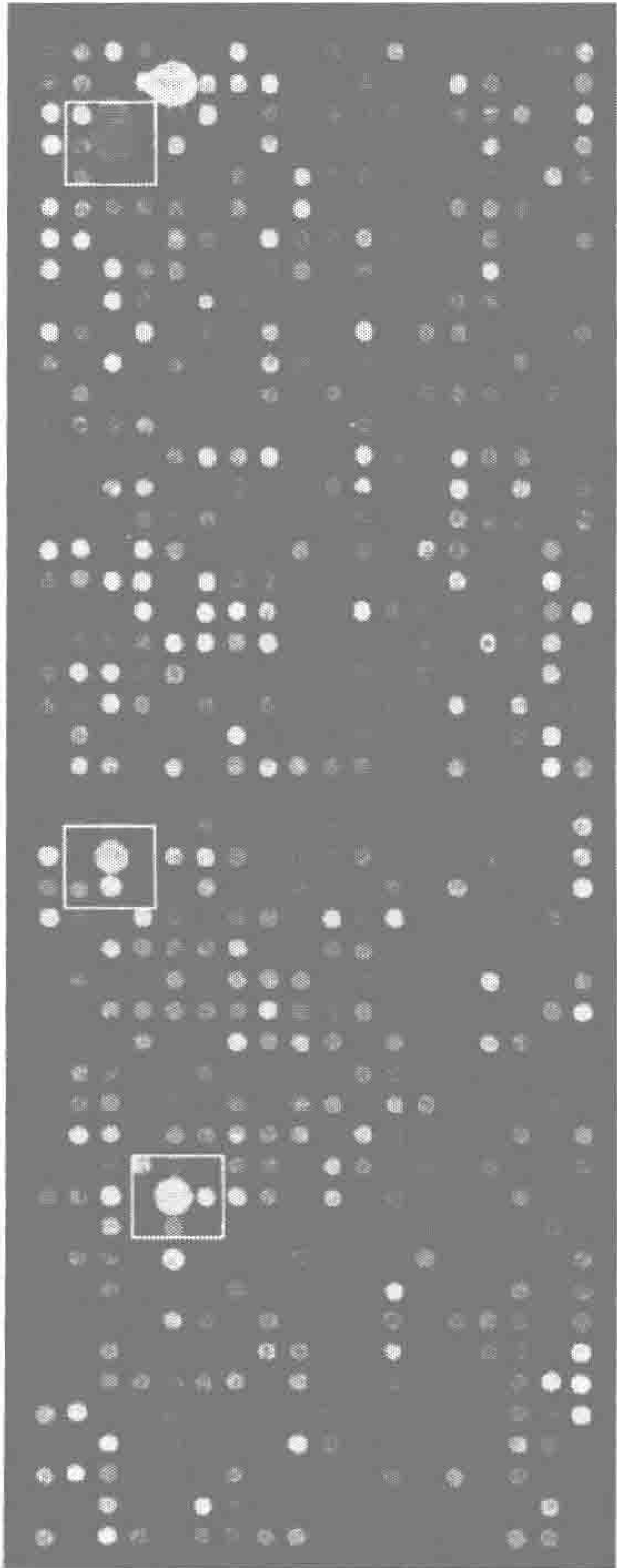
可能的解决方法

- 检查加样板中 DNA 的容量。在 384 孔板中，不应超过 1 mm 或者大约 8 μl。
- 检查和清洁点样针。
- 检查和修改点样仪性能设置中有关蘸墨、停留和加速时间的项目。
- 使用建议的点样针/基底组合。检查玻片的批次。
- 相对湿度应该保持在 50%～60%。

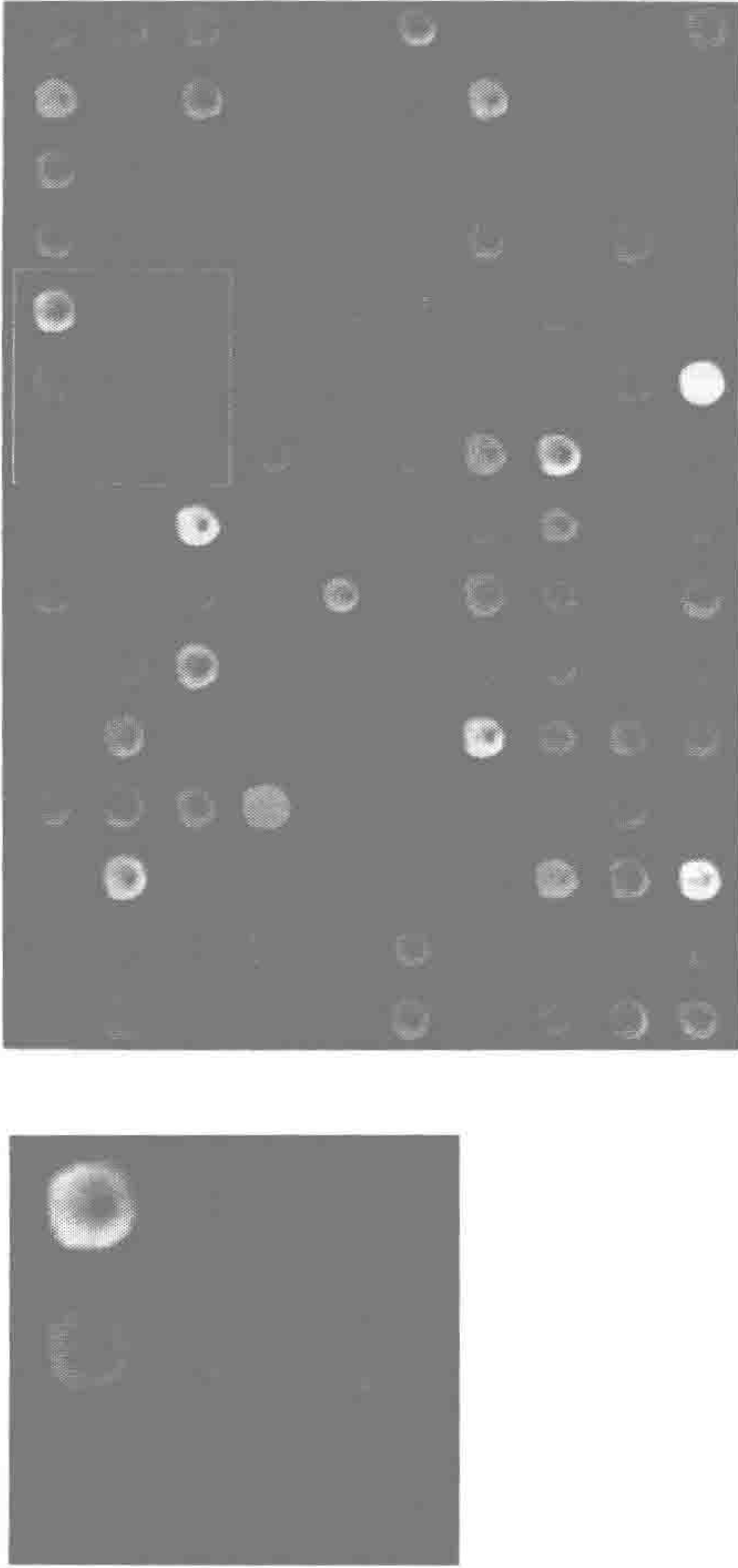
图像 6：圆环样 **246**

问题

样点中 DNA 的分布不规则。DNA 的沉积包含了点样针尖的几何形状、基底的疏水特性、点样溶液的润湿特性、点样过程中的湿度、接触过程中的向下作用力和点样针尖离开时向上的加速这些因素之间的相互作用。上图的样点中 DNA 的不规则沉积，呈环状或"油炸圈饼"图案。下图是上图中方框区域的局部放大。



图像 5



图像 6

原因

- 点样针、载玻片和（或）点样缓冲液组分的不适当组合。
- 点样仪的设置不适当。
- 玻片干燥太快。

可能的解决方法

- 尝试不同的载玻片基底，要注意具有不同疏水特性的基底。
- 比较不同的点样针。
- 在点样溶液中加入 DMSO 或甜菜碱；
- 在 50%～60%的湿度条件下进行点样。
- 对点样头的接触速度和向上的加速过程进行实验比较。
- 点样后对载玻片进行重新水化（见第 2 章）。

247

图像 7：边缘干燥

问题

由于杂交过程中的脱水导致主要在玻片的边缘产生严重的杂交背景。盖玻片难移去也表明了干燥过度。

原因

- 杂交过程中未能提供足够的湿度。
- 对于在 65℃下进行的杂交反应，该问题更加严重。
- 杂交时间过长。

可能的解决方法

- 在杂交室内放入湿润的 Whatman 滤纸。
- 考虑使用不同的杂交室设备
- 考虑使用加入甲酰胺的杂交液，并在 42℃下孵育。
- 用空白的载玻片进行实验，以确保达到正确的湿度平衡，并且没有干燥/凝结。
- 杂交反应进行约 16 h。

图像 8：边缘褪色

问题

在微阵列的边缘部分图像逐渐褪色。



图像 7



图像 8

原因

- 建立杂交反应时靶分子的不均匀分布。
- 探针与靶分子的接触受到扩散限制。
- 盖玻片边缘的凝液稀释了靶分子。

可能的解决方法

- 按照正确的方法建立杂交反应。
- 考虑改变杂交体积。
- 考虑使用一种可以在杂交过程中混合靶分子的杂交工作站。
- 确保合适的湿度，不要在玻片表面形成凝液。

248

图像 9：高背景：荧光

问题

上图为高强度的不规则背景。下图有严重的背景问题，图案表明在最终的离心过程中，背景流过了整块玻片的表面。

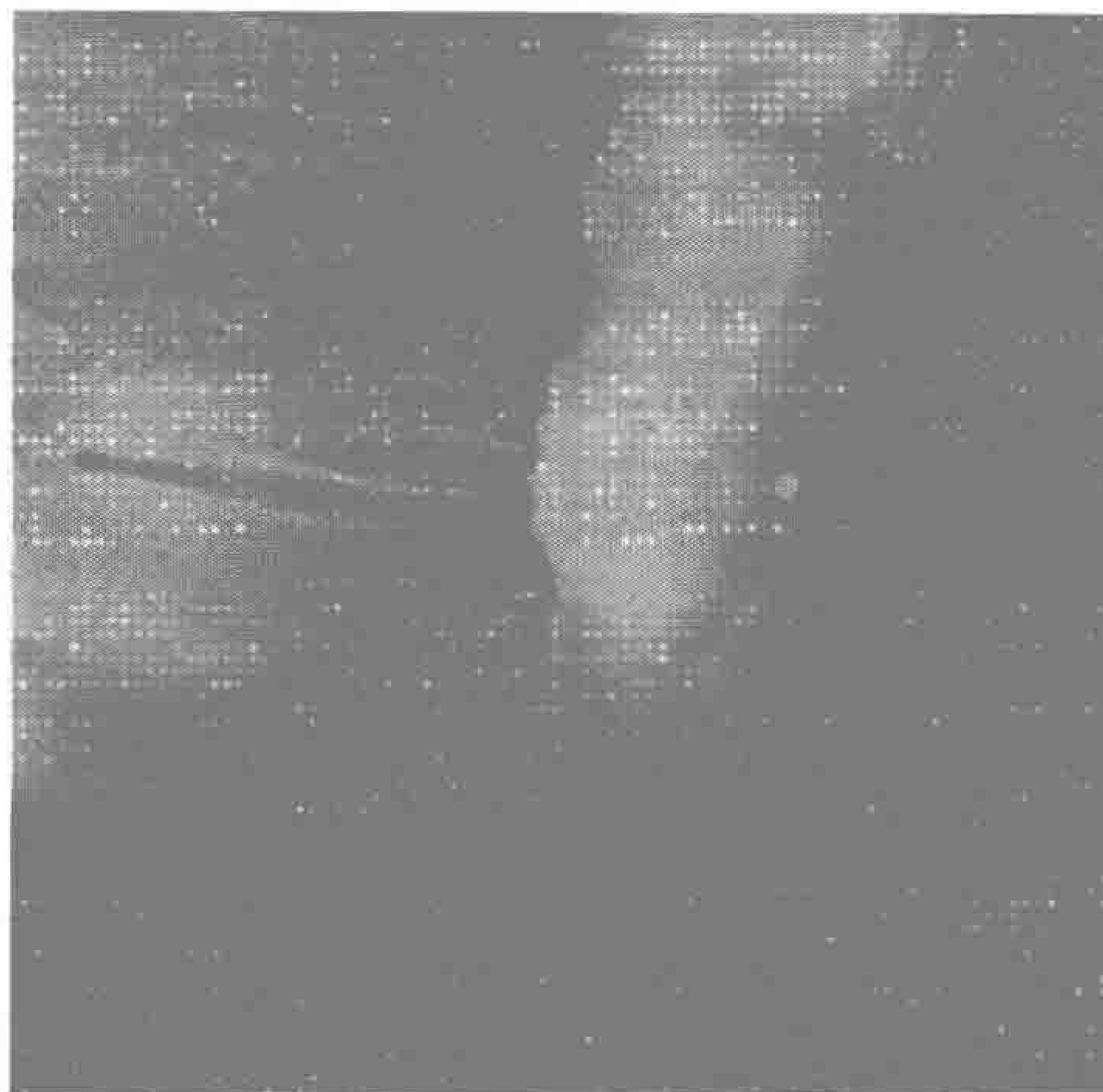
原因

- 载玻片基底本身固有的高强度背景荧光。
- 未能从靶分子中充分地除去未掺入的核苷酸。
- 在进行杂交前没有加热靶分子。
- 杂交溶液具有高强度荧光，清洗不充分。
- 在首次清洗时，盖玻片未能从载玻片上移去。
- 在清洗步骤结束时，未能除去残留的溶液(特别是 SDS)。
- 在清洗过程中发生脱水。
- 来自皮肤的油脂或者是手套上的乳胶。
- 制备靶分子使用的乙醇产生的荧光。

可能的解决方法

- 在点样之前通过扫描对新批次的玻片进行检查。
- 对靶分子进行足够的纯化。
- 在马上开始杂交之前将靶分子在 100℃加热 2 min。
- 使用大量的溶液和大容器进行清洗，避免在不更换缓冲液情况下洗涤多块玻片。
- 按照推荐的次数进行清洗，并作振荡。
- 确保盖玻片脱落。
- 避免在清洗缓冲液中使用过多的 SDS，并确保清洗中全部的 SDS 都被去除。
- 清洗过程中避免载玻片干燥。操作要迅速，在转移载玻片之前要准备好容器。





图像 9

- 清洗结束后立即离心载玻片。不要长时间放置在架上和分批收集。
- 戴无粉的丁腈手套。
- 更换乙醇的品牌。使用 Gold Seal 或者 Warner Graham 出品的乙醇。

249 图像 10：高背景：黑洞

问题

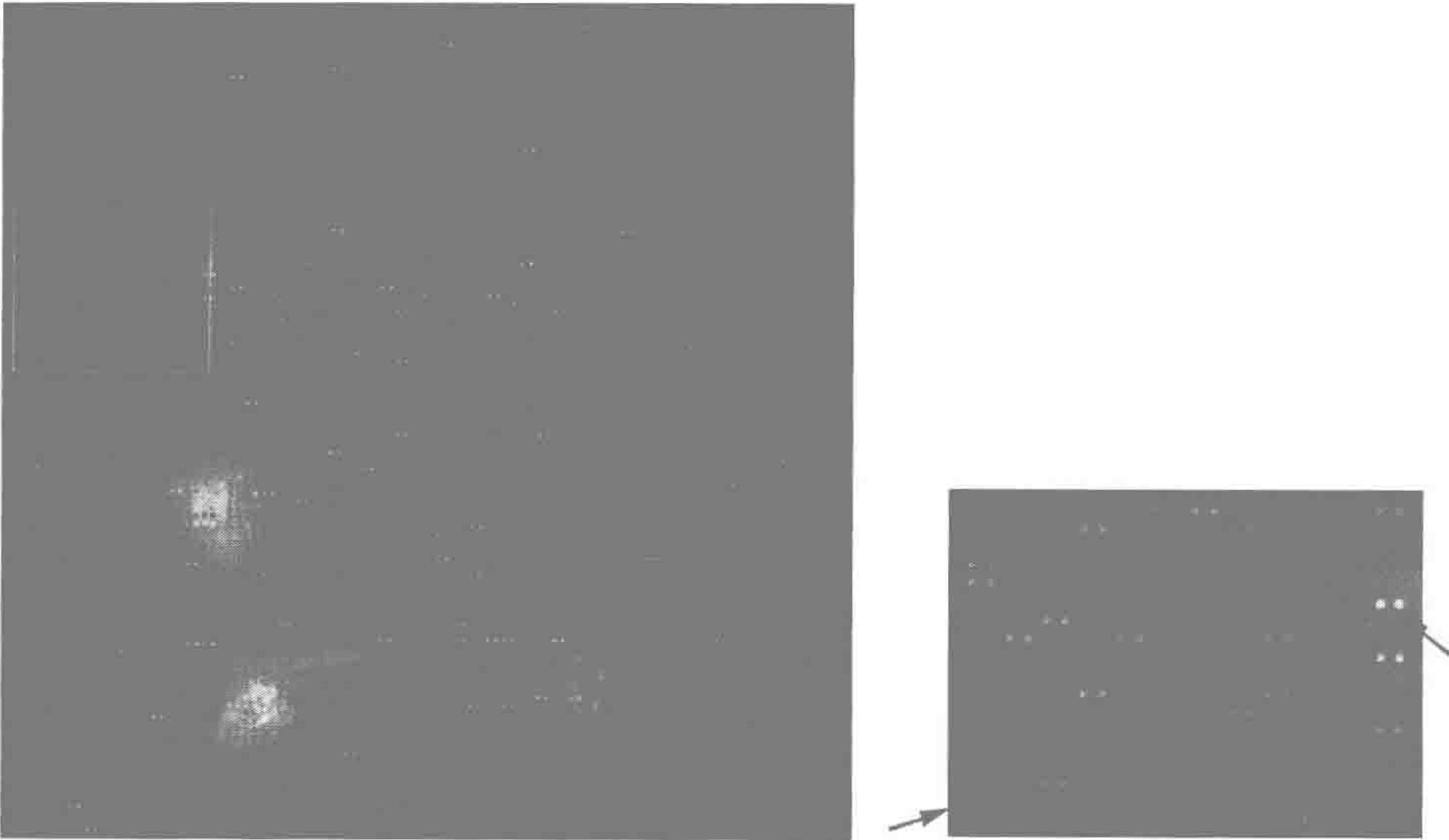
信号逆转：黑洞。左图中标记的靶分子和基底的结合呈现出高背景，其中许多样点出现“黑洞”，也就是说，样点的荧光强度要比周围的背景还低。右图为左图中方框区域的局部放大，显示出点样区域存在的 DNA 起着封闭剂的作用，阻止了靶分子的结合（左侧箭头），不能为杂交反应提供靶分子导致样点上的信号微弱，除非是高度表达的基因（右侧箭头）。

原因

- 点样后未能充分封闭载玻片。
- 杂交溶液中省略了封闭试剂。
- 靶分子制备得不正确。

可能的解决方法

- 确保进行了正确的封闭反应。封闭溶液在使用前即刻新鲜配制。
- 检查杂交溶液，要与载玻片基底相匹配。
- 按照推荐的方法制备靶分子，包括清除步骤。
- 考虑使用封闭要求严格度较低的载玻片基底。



图像 10

250 图像 11：不规则样点形态

问题

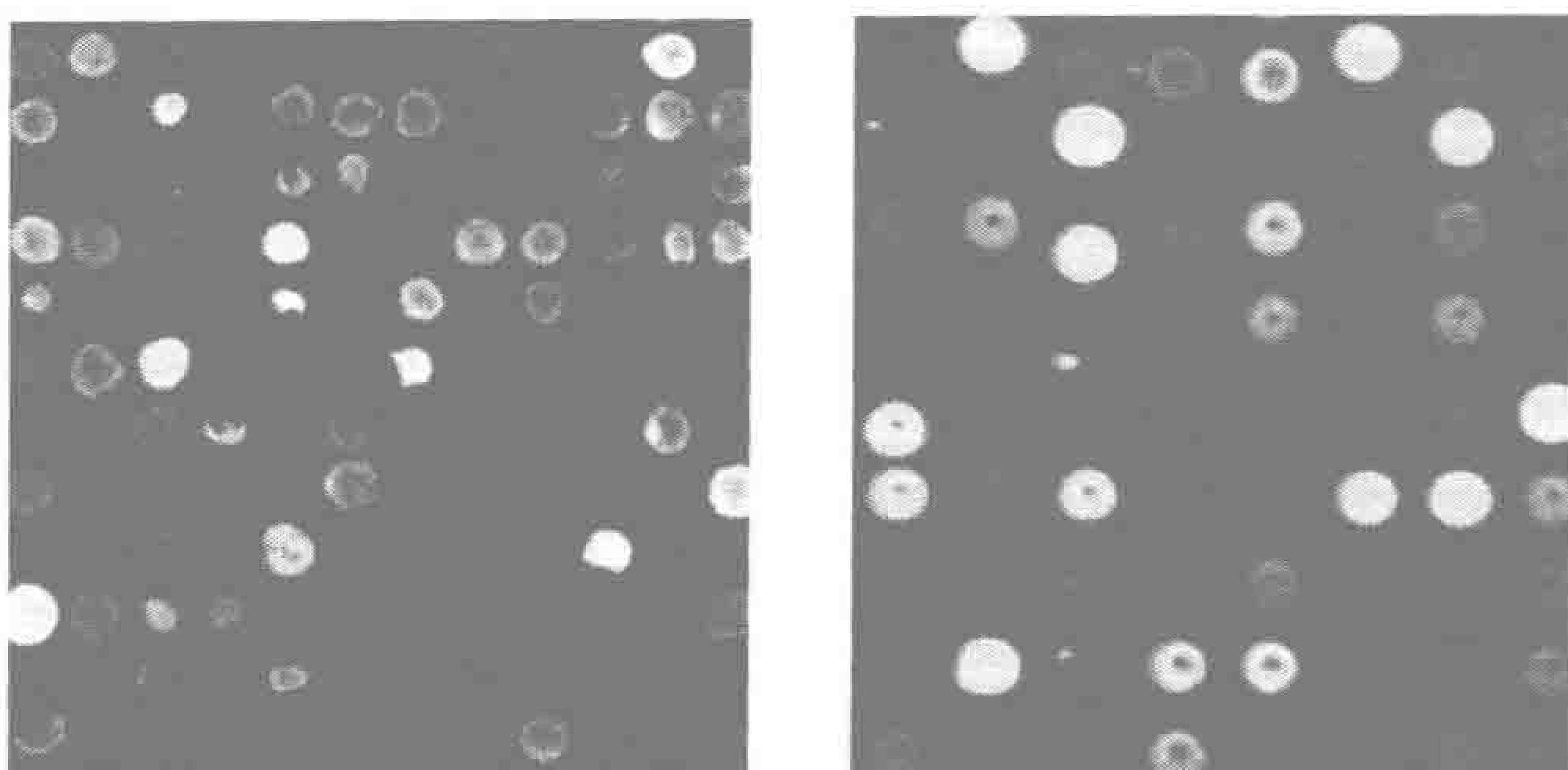
由于点样针的缺陷导致不规则样点形态或者是不规则的布局。上图中由于针尖设计上的不足或者针尖损坏而导致样点在形状上不均匀。下图是由于构成点样头的其中一根针有轻微弯曲而造成点与点之间的距离差错。如这里所示，如果弯曲的针能够在其针室中自由的转动，样点的位置会发生相当大的差错。

原因

- 点样针设计有不足。
- 点样针发生弯曲。
- 针尖损坏。

可能的解决方法

- 新的点样针。针由于太细而无法试图重新加工。
- 操作点样针时要格外小心，特别是从点样头中取下、更换时，或者超声处理时。



图像 11

图像 12：信号强度低

问题

信号强度低。尽管检测到许多的点，但是动态范围有限，并且大部分信号只比背景高数倍。许多样点的低水平杂交预示着这是一种非特异性的信号。



图像 12

原因

- 点样材料的纯度或者是浓度不够。
- 稀释的杂交条件。
- RNA 的数量不够。
- 标记不够。
- 扫描仪的故障。

可能的解决方法

系统地进行整个过程操作，确保在每一步骤都包含了适当的对照。按照推荐的方法进行以下的操作：

- PCR 的定量和清除。
- RNA 的纯化和标记。
- 检查染料和酶的批次，在凝胶上检查探针的质量。
- 更换杂交试剂。
- 检查扫描仪的性能。激光器的故障可能是问题所在。
- 与设备的其他使用者进行经验对比。
- 保留一个 QC（译者注：quality control，质量控制）参考集：保留一套载玻片、RNA 和试剂的证据集合用于解决问题。

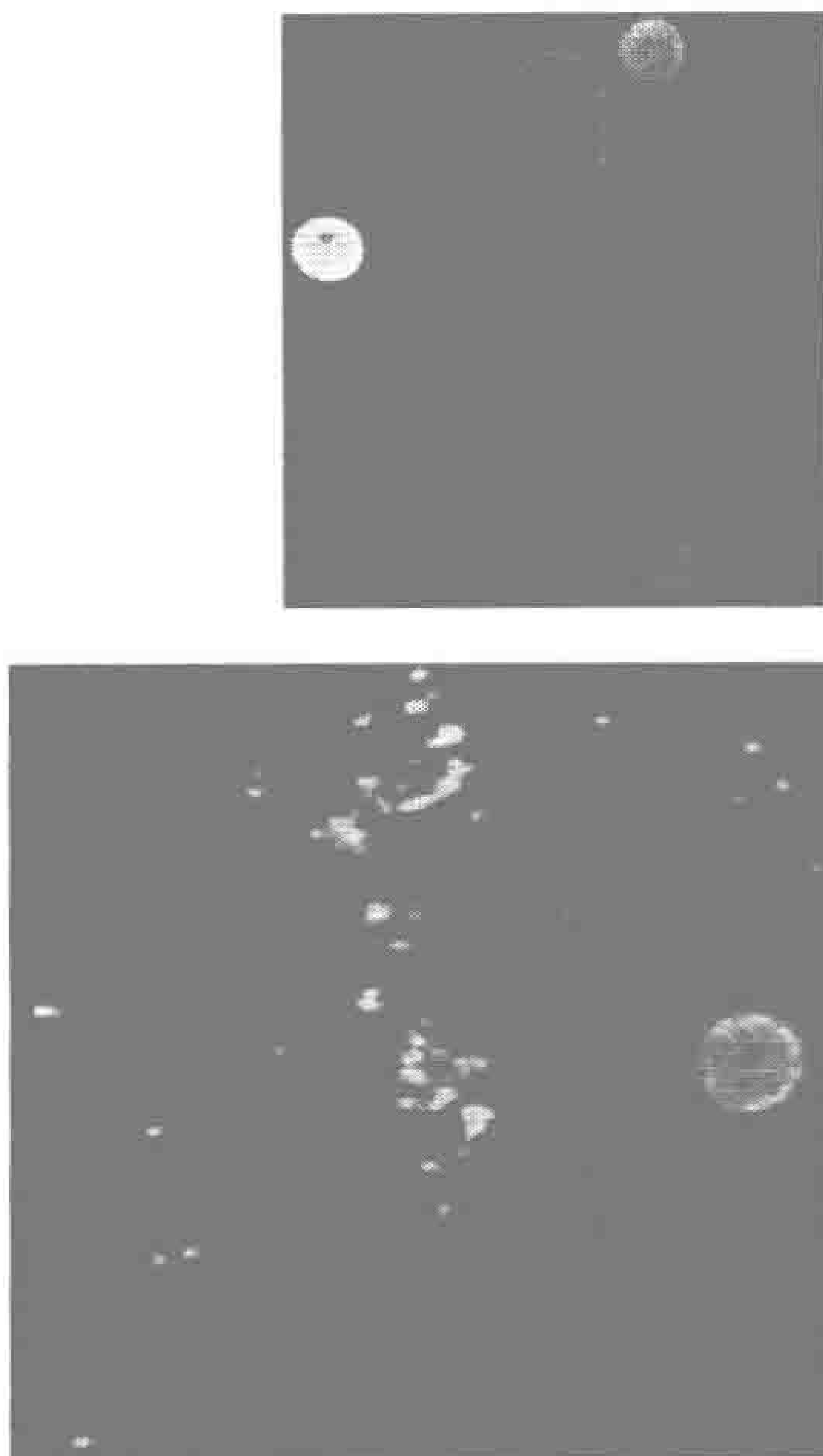
图像 13：颗粒污染

问题

干扰对样点的定量。上图是受到纤维污染。下图是受到粉末或者来自靶分子中的颗粒的污染。

原因

- 来自工作环境中的颗粒物质（工作台垫、手套、棉纸、平常的工作环境——译者注：指未进行净化处理的环境）。
- 受到污染的载玻片或者盖玻片。



图像 13

• 靶分子清除步骤中未被离心去除的树脂。

• 清洗缓冲液中存在的颗粒。

可能的解决方法

• 准备杂交时使用铝箔，或者将工作台垫有塑料的一面向上放置。

• 使用无粉的丁腈手套。

• 杂交前使用不含有油的压缩 N_2 除去载玻片和盖玻片上的灰尘。

• 使用为微阵列而设计的盖玻片，如那些具有保护性背衬的盖玻片；

• 在进行玻片杂交前，将靶分子溶液以 10 000 g 离心 5 min。不要碰到离心后的小团块。

• 使用前过滤杂交溶液和清洗缓冲液。

图像 14：点样针堵塞

问题

点样针堵塞（鹅毛笔针）。在所示部位（箭头）由于完全（上图）或部分（下图）堵塞而影响到探针的沉积。

原因

• 点样材料的制备有缺陷。

• 点样过程中点样针的清洁不充分。

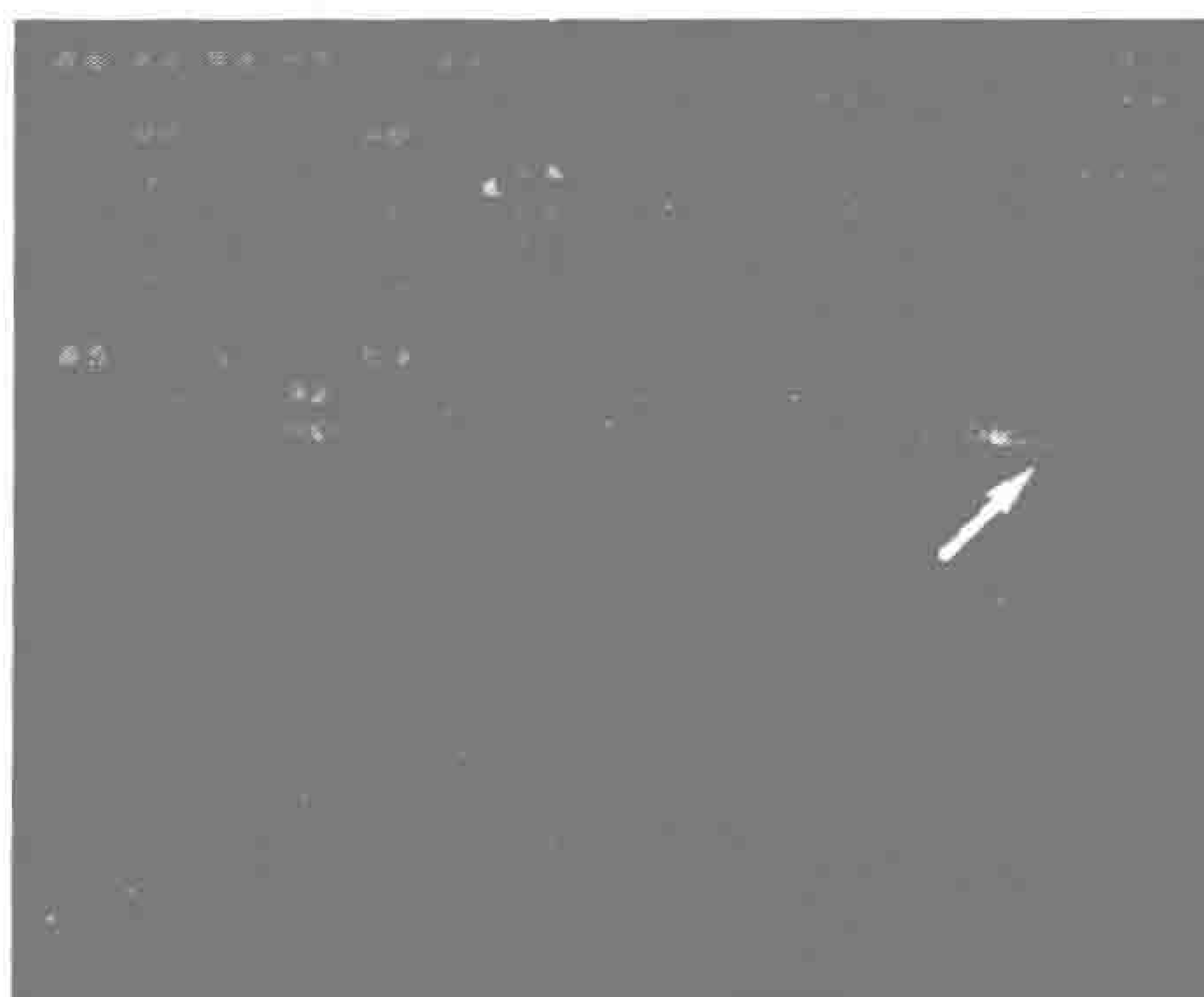
• 点样针在设计上有缺陷。

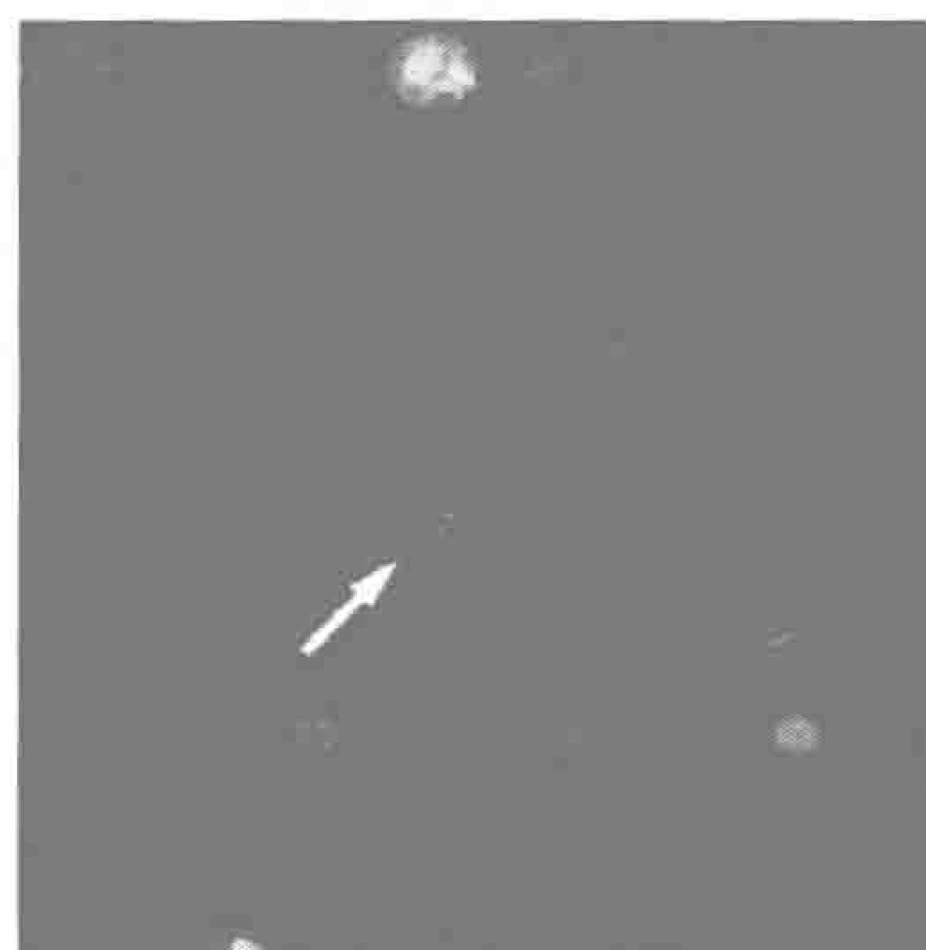
可能的解决方法

• 制备点样材料（见第 1 章，方案 1～3）时要特别注意除去颗粒和蛋白质材料。

• 检查清洗工作站的工作情况。最好使用带有超声波洗浴的清洗工作站。

• 按照第 2 章，方案 4 中介绍的方法清洗点样针。





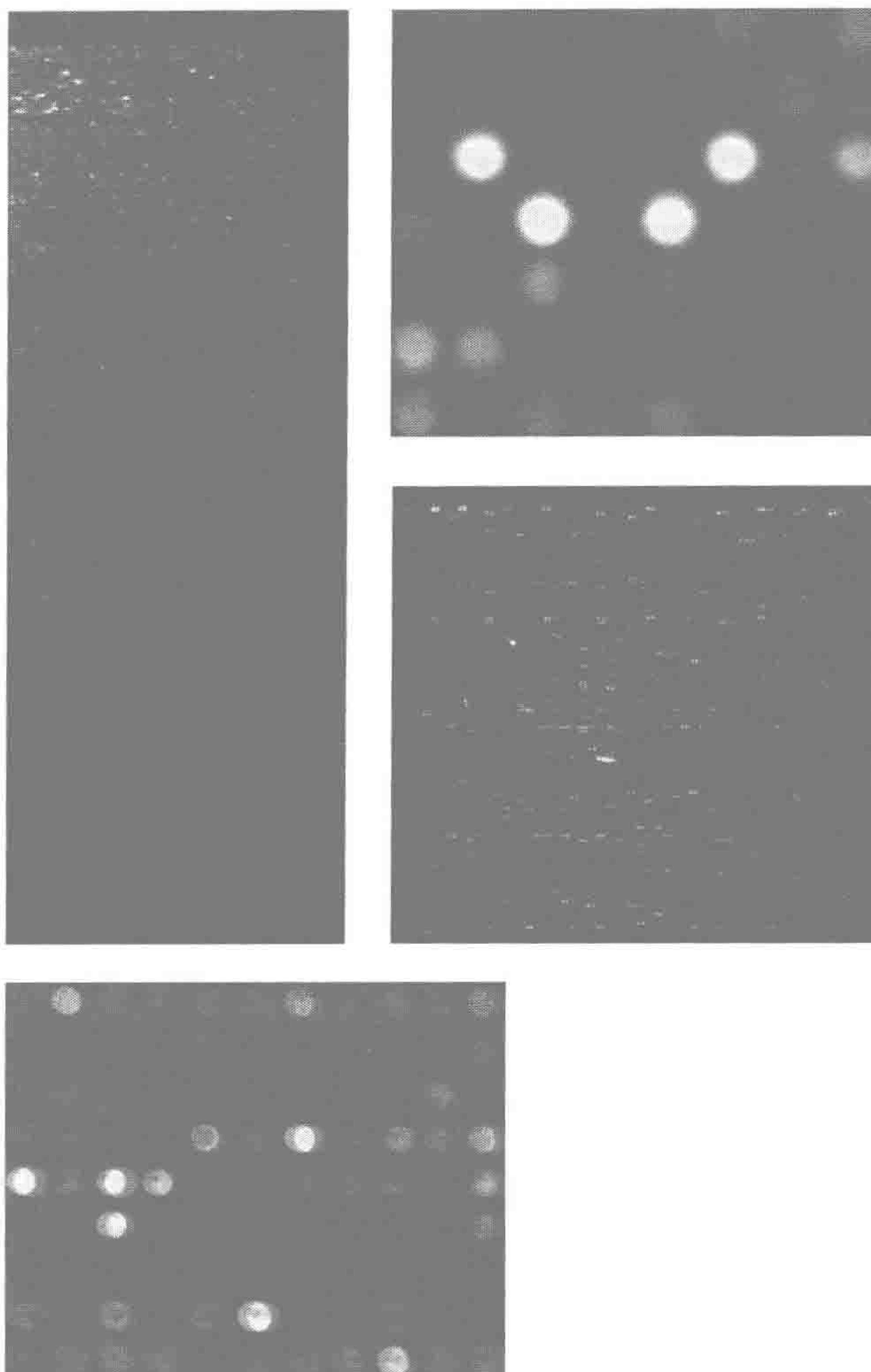
图像 14

252

图像 15：扫描仪故障

问题

载玻片在扫描仪中的位置：（左上图）信号沿着载玻片的轴向减弱，因为载玻片在扫描仪的托盘上没有放平。（右上图）样点的形态良好，但是载玻片焦距没有对准。扫描仪光学系统故障。（右下图）沿着载玻片横向规则的强和弱信号图案。（左下图）样点聚焦相对较好，但是红色和绿色通道没有对准。



图像 15

图像 16：日复一日点样中的差异

253

该图中测量了微阵列点样的天数对信号强度的作图。使用带有 16 根点样针的印刷头在 5 天内制备微阵列，每根针点样一个 19×21 的大方格。玻片与相同的 RNA 杂交，由 Cy3 或 Cy5 标记。x 轴对应于在 5 天的点样中逐步集中起来的 19×21 格子中的每一个坐标。y 轴表示每组 16 个点的平均强度，也就是说，代表了给定的坐标位置。(D) 表示了 Cy3 通道的强度（$\log_2 G$）；（C）表示 Cy5 通道的强度（$\log_2 R$），覆盖图（B）是强度总和的平均值（0.5×）（$\log_2 R + \log_2 G$）以及每个通道的相对强度 (A)$\log_2$（R/G）。可以看出样品之间信号强度有着相当大的变化，与点样的日期相符。

问题

信号输出强度的变化是点样时间的函数。

原因

引起这种变化的原因尚不清楚，但是它们可能与湿度，点样针的性能差异有关，或者是依赖于操作者的因素。

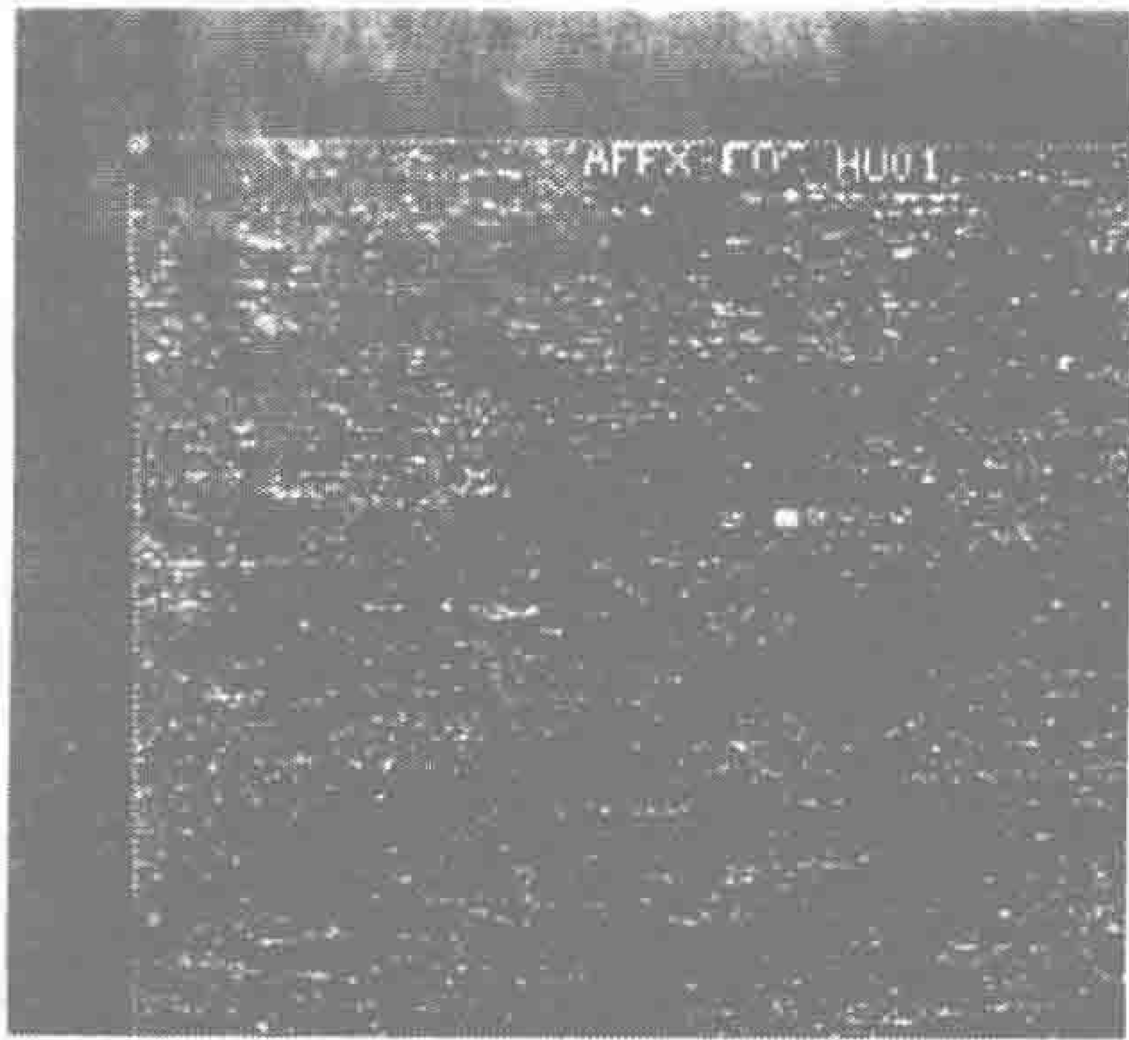
解决方法

由于引起问题的原因不明，因而纠正方法不明显。研究结果强调了一个事实，即必须考虑到不同

实验之间依赖于点样的差异来源。

图像 16

254



图像 17

图像 17：高度不规则的背景

问题

高水平的不规则弥散背景。

原因

芯片外表面有油脂或者灰尘，或在内表面有黏合胶。

解决方法

检查以确保玻璃表面干净，然后重新扫描。如果问题仍然存在，样品必须重新杂交。

图像 18：颗粒污染

问题

细的颗粒物质。颗粒背景低，可以接受（左图）或

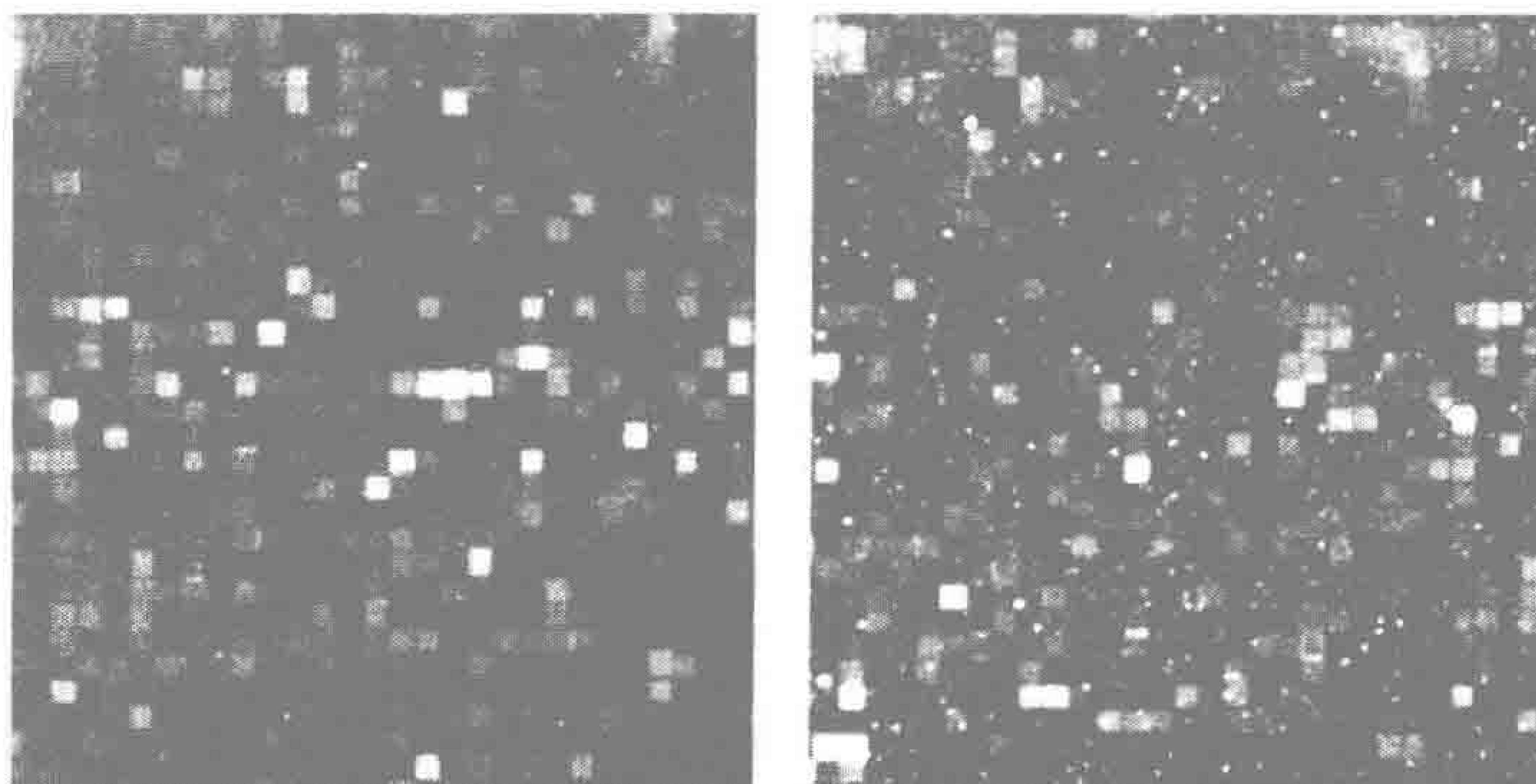
者情况严重需要重新实验（右图）。

原因

杂交溶液/样品溶液中含有盐、蛋白质或颗粒物质。

解决方法

使用0.2 μm过滤器对杂交样品溶液进行预先过滤。如问题仍然存在，可以从Affymetrix的技术支持获得包含了附加封闭试剂的缓冲液配方。



图像 18

图像 19：明亮的斑块

问题

明亮的斑块/条纹。高水平的局部背景呈现出条纹状（左图）或者斑块状（右图）（见箭头）。

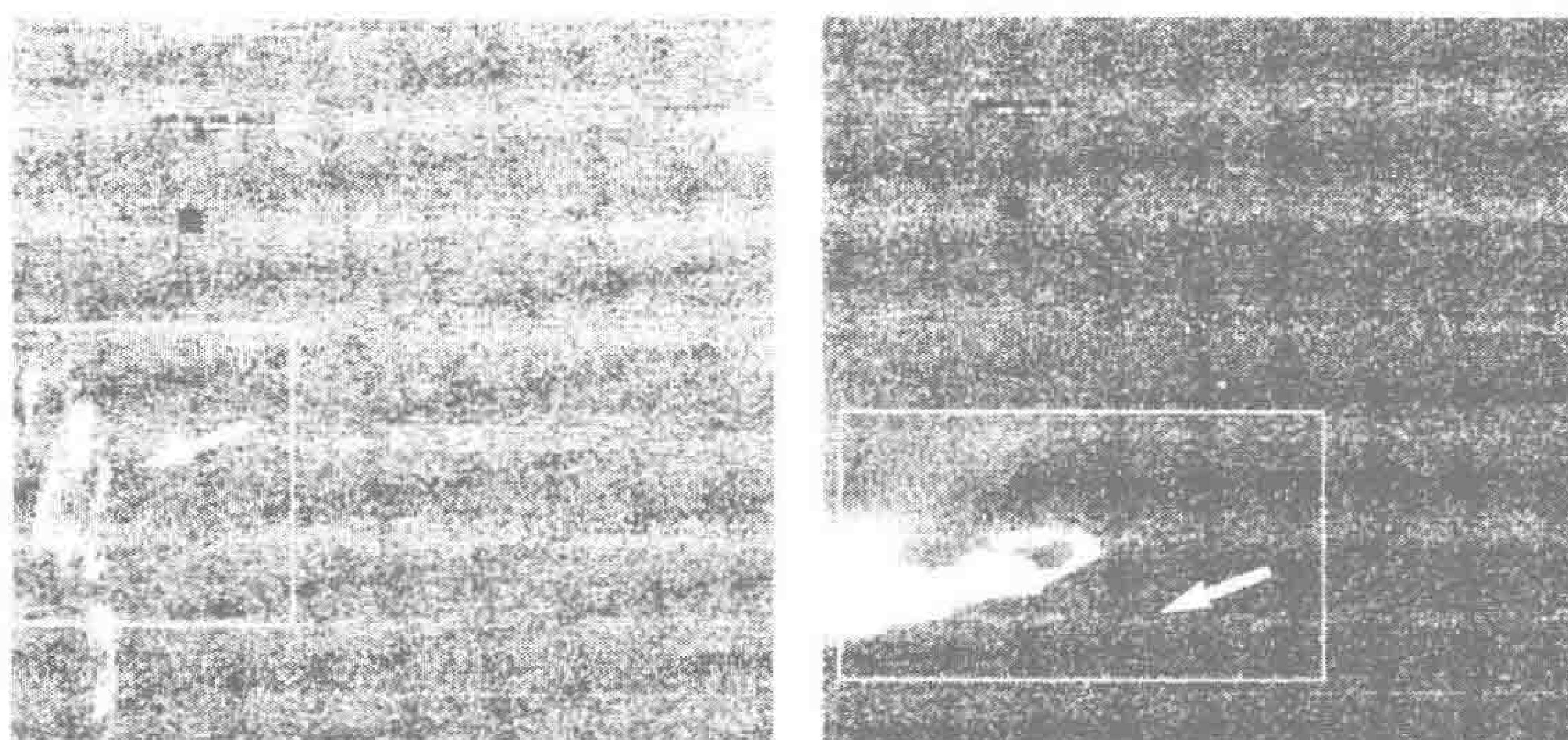
原因

黏合胶和链亲和素——藻红素的非特异性吸附。

解决方法

如果瑕疵覆盖了芯片上相当大的部分，那么样品必须重新杂交，如果有必要，弃去该芯片。

其中一些图像中出现的有规律强度变化条纹，事实上并不出现在数据中，而是由于将原始的光栅图像高度压缩为可打印的像素密度而导致的图像取样假象。



图像 19

图像 20：非特异性信号

问题

信号普遍较弱而不能区分特征。

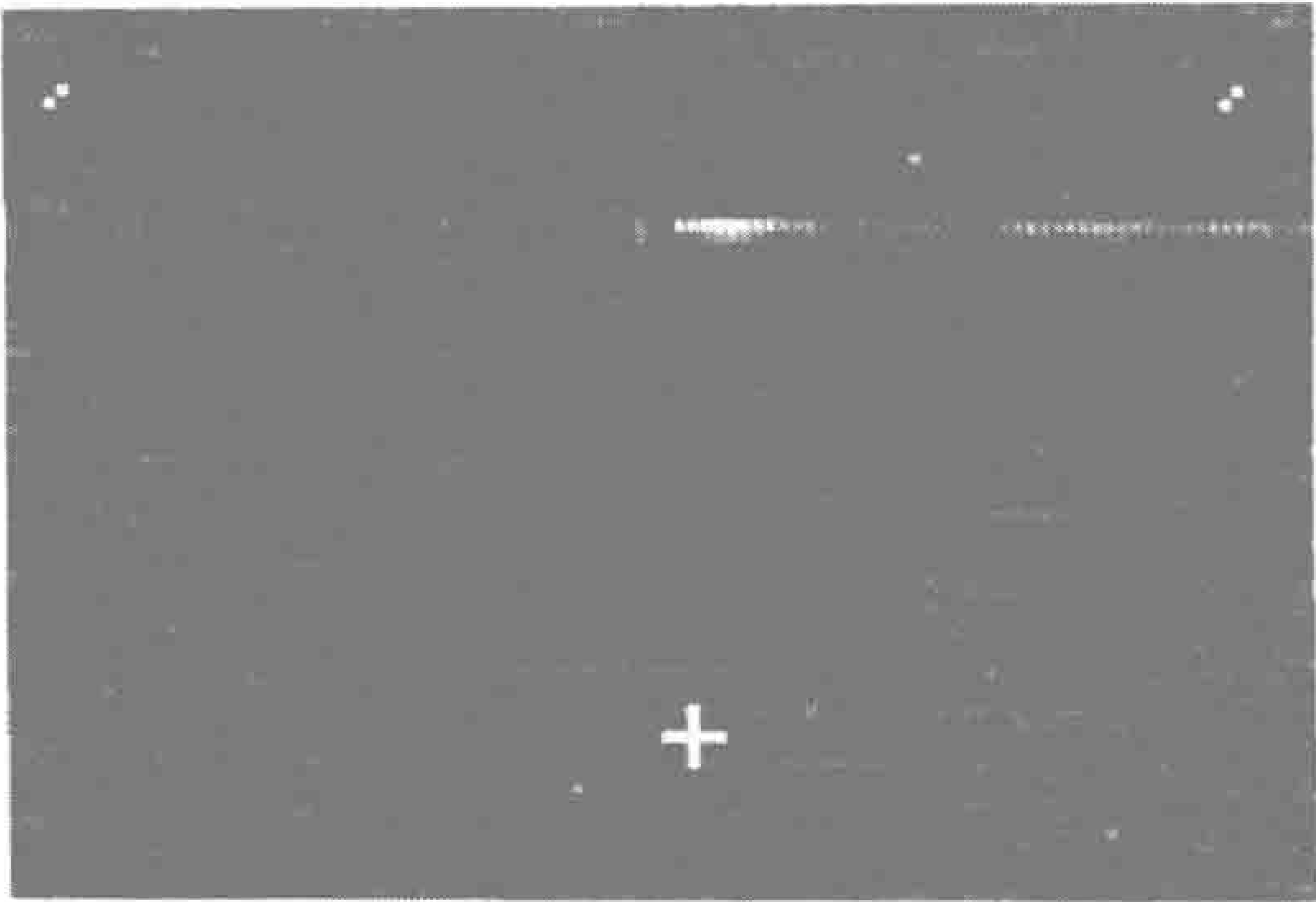
原因

靶分子的质量差；用于突变分析的杂交或清洗缓冲液温度不恰当、放置过久或染色后的制备物发

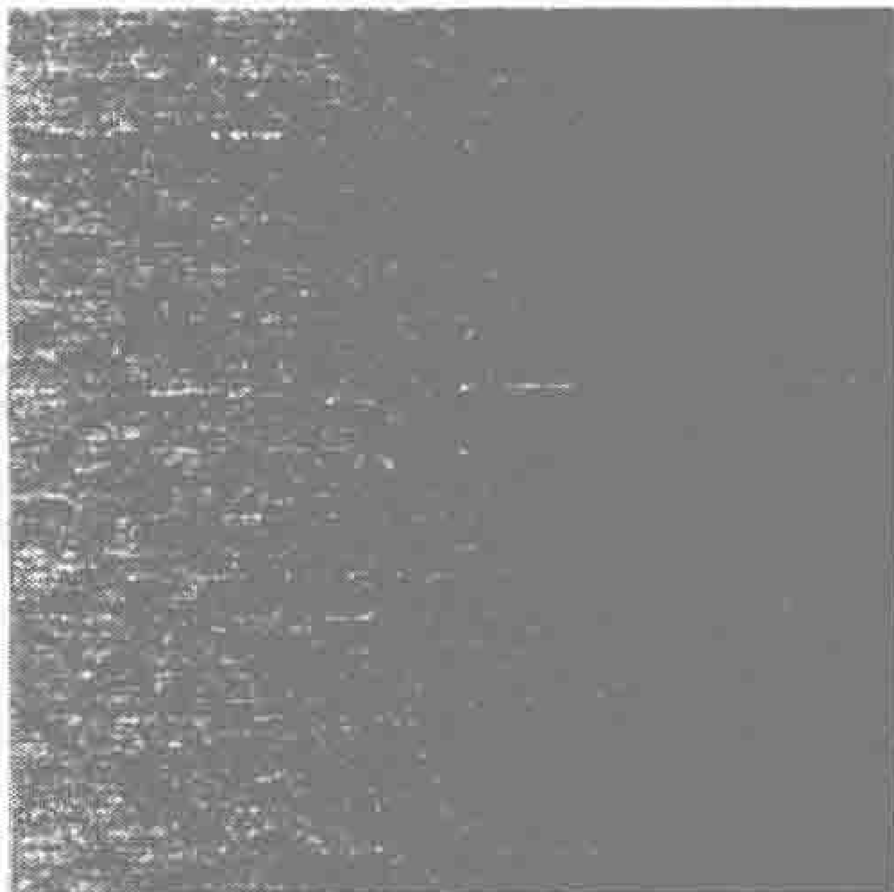
生漂白。

解决方法

在其他使用相同批次芯片的杂交中检查该问题。使用先前成功的杂交样品测试来自该批次的其他芯片。如果问题出在芯片合成上，使用不同批次的芯片重复杂交反应；如果靶分子有问题，用起始的RNA重新处理靶分子并杂交。检查杂交箱的温度。检查自动洗涤程序的脚本。比较来自清洗工作站不同清洗模块的结果。



图像 20



图像 21

图像 21：扫描仪故障：信号缺失

问题

沿着芯片的表面，信号逐渐消失。

原因

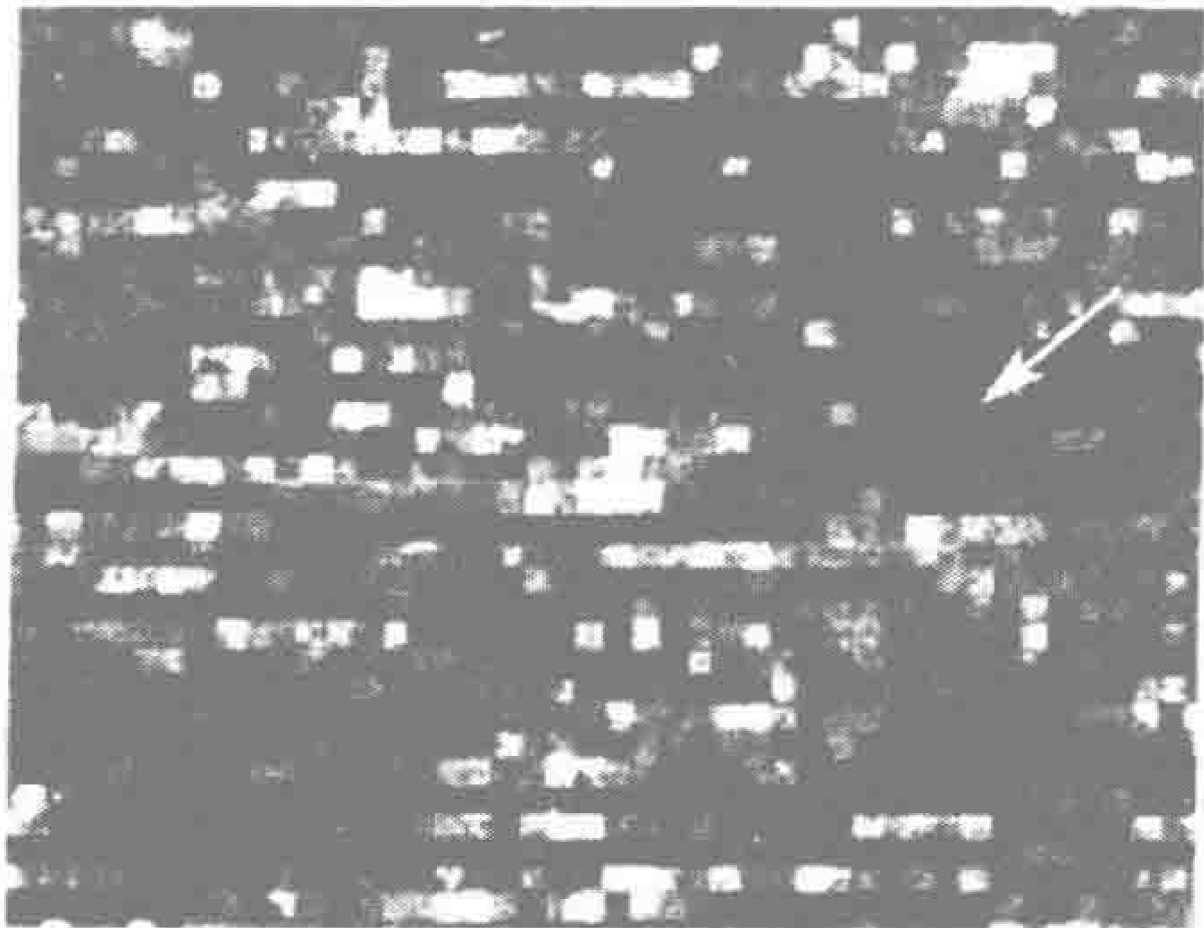
扫描仪焦距没有对准，清洗或者染色不均匀。

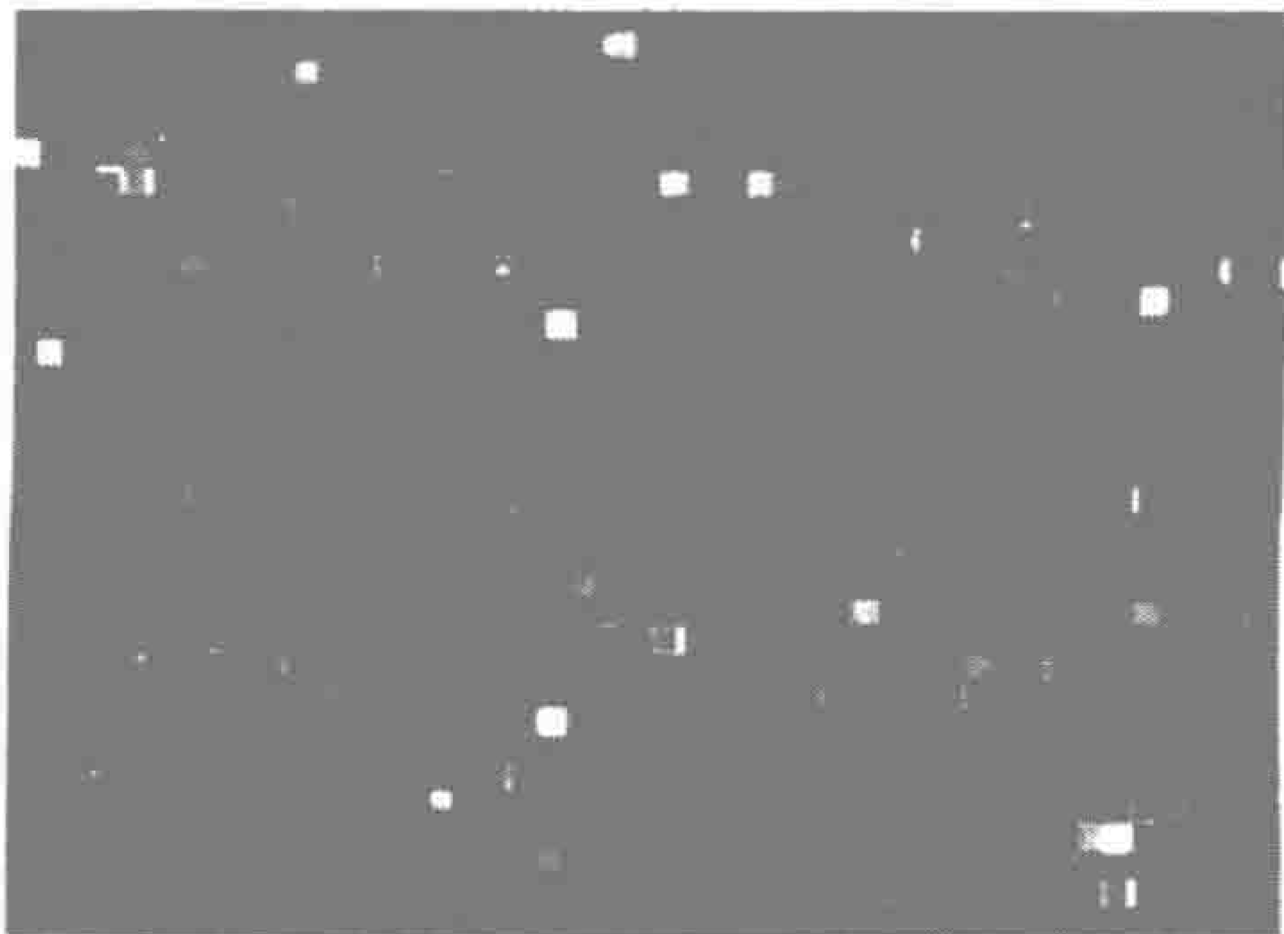
解决方法

重新扫描芯片（如果有可能的话在另一台扫描仪上）可能解决聚焦问题。否则，样品必须在另一块芯片上重新杂交。

图像 22：芯片缺陷

芯片制备过程中可能的缺陷。左图表示样点上的缺陷，内部有洞（箭头）。缺陷可能是由老版本的 Affymetrix 方案中的防沫剂引起的，不包含该组分的较新的配方不容易产生此缺陷。右图表示样点带有明亮的边缘，可能是与 cDNA 断裂有关的问题。





图像 22

图像 23：图像表面刮痕 256

问题

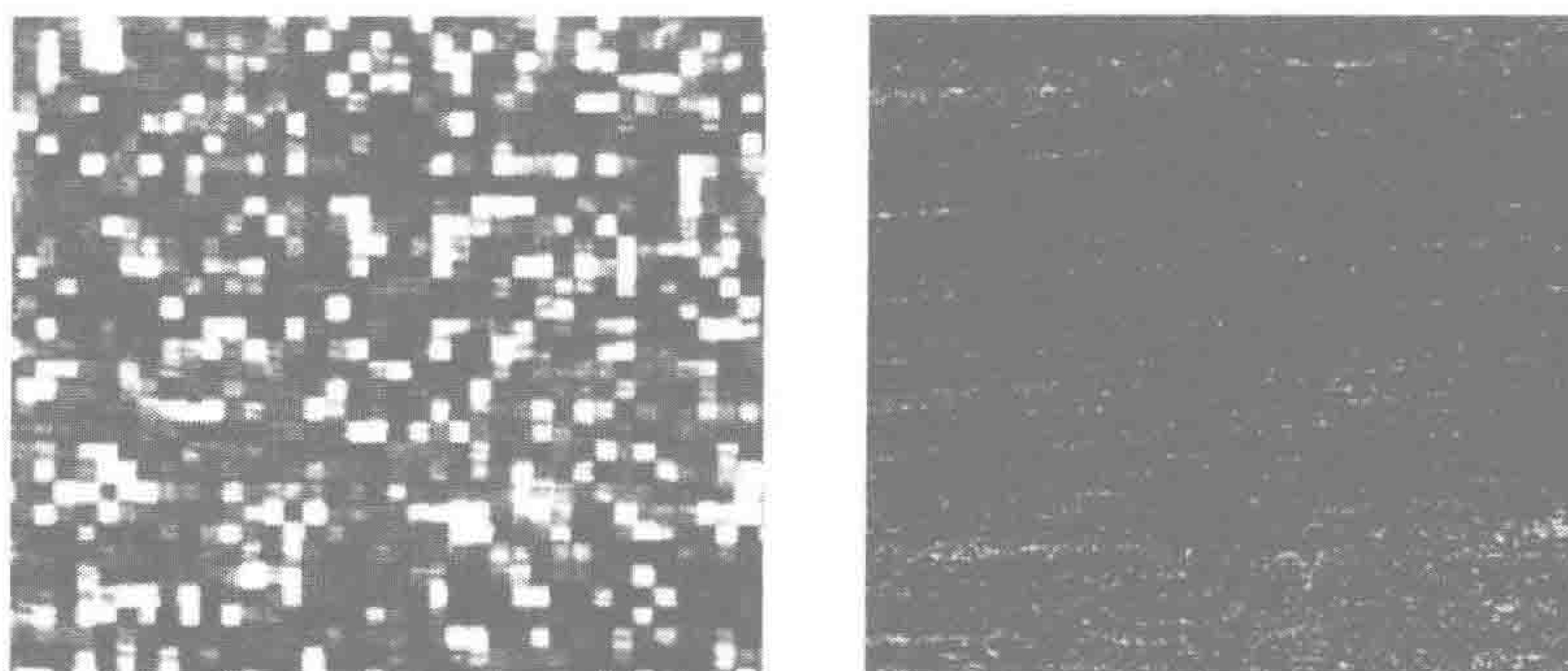
芯片表面受到机械损伤。上图表示有细微的刮痕。下图表示有严重的刮痕。

原因

刮痕可能是在芯片封装或者杂交过程中产生的。刮痕是由从流体注入端口挖取的隔膜碎芯或其他未经证实的颗粒造成的。

解决方法

将靶分子与新的芯片重新杂交。



图像 23

信息栏：Alexa 染料 257

Alexa 染料是由 Molecular Probes 公司合成和销售的一系列特有的七种荧光化合物（www. molecularprobes. com）。染料的名称（Alexa 350，430，488，532，546，594，658）表示了它们大约的最大激发波长（图 3-24，图 3-25）。通过磺化香豆素或罗丹明环上的氢原子而获得这些染料，已知这个过程可以在水相媒介中增加其他许多荧光化合物的亮度（Wessenndorf and Bralje 1992；Mujumdar et al. 1993）。与磺化的青色素染料 Cy3 和 Cy5 相比，Alexa 染料对较宽的 pH 范围是不敏感的，并且更加抗光漂白。除此之外，他们的琥珀酰亚胺酯可溶于水，可以在没有有机溶剂的情况下进行染料的修饰结合（Panchuk-Voloshina et al. 1999）。

Alexa 546 是一种橙色染料，可以获得琥珀酰亚胺酯的形式，或者是结合到 dUTP（546-14-dUTP，Molecular Probes，Inc）。和其他罗丹明染料一样，它具有高的摩尔消光系数（ε=104 000，见关于**测量 DNA 的染料标记效率**信息栏）。在面对面的直接比较中，Alexa 546-14- dUTP 标记的第一链 cDNA 在 DNA 微阵列中发射的荧光要比 Cy3-dUTP 标记的 cDNA 强大约三倍（Wildsmith et al. 2001）。尽管在使用标记的第二链 cDNA 的杂交实验中，两种染料给出了相当的结果（图 3-25），但使用 Alexa 546 标记效率要高一些（L. Rose 未发表结果）。未掺入的 Alexa 546 和 Cy3 一样，可以按照方案 19 中的介绍，从使用过的标记反应中回收。

在本书写作时，只有一种 Alexa 染料的核苷酸结合物可以购买到（546-14-dUTP）。随着更多的结合物投向市场，Alexa 染料将会给在 DNA 微阵列实验中对靶分子标记占

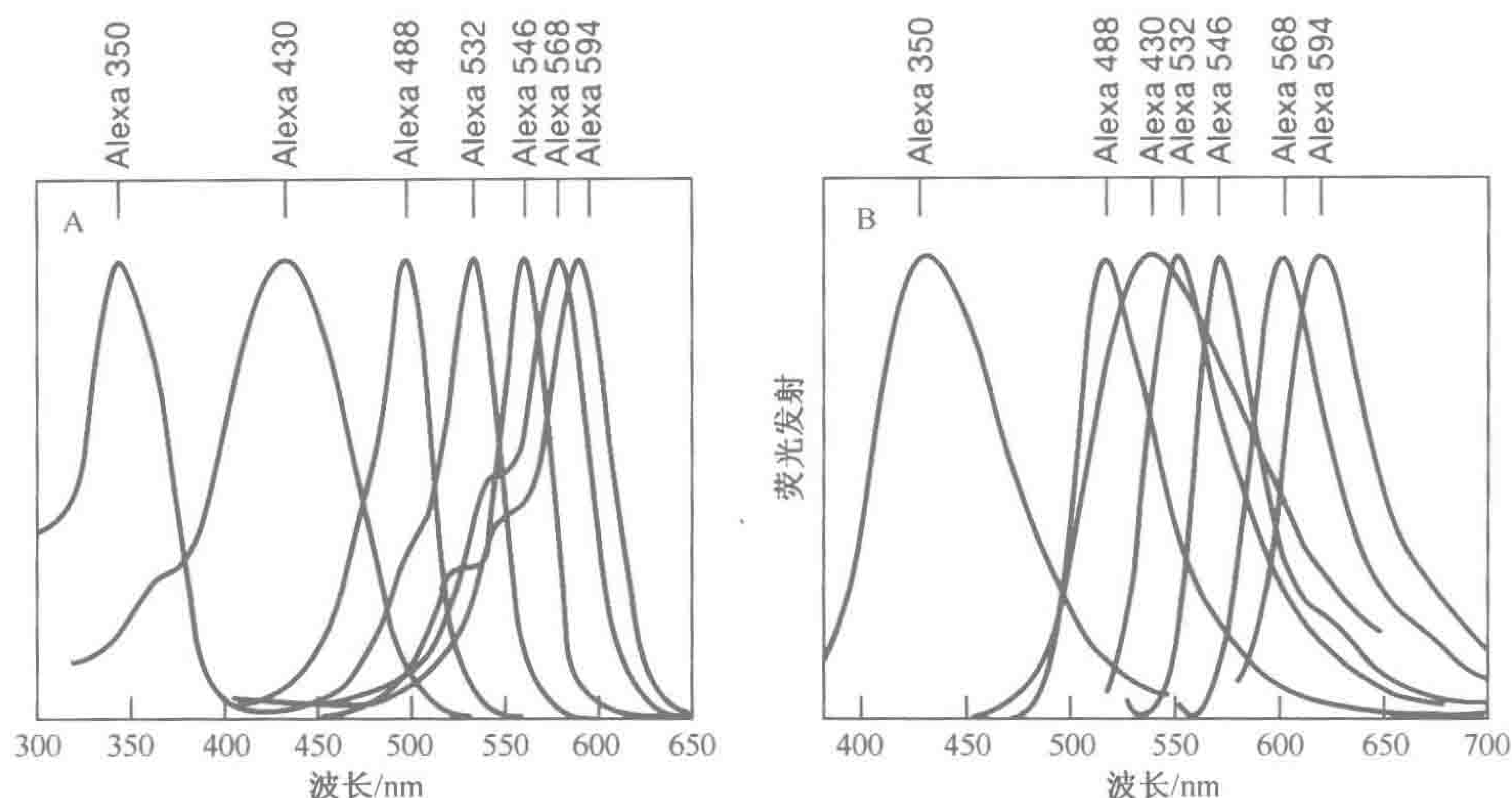


图 3-24　结合到羊抗鼠抗体上的 Alexa 染料的归一化吸收（A）和发射（B）光谱

除了最大发射/吸收峰红移数纳米以外，结合物的光谱形状与游离的 Alexa 染料非常相似

（经许可重新绘制，来自 Panchuk-Voloshina et al. 1999）

统治地位的磺化青色素染料带来挑战，特别是如果它们能保持现在的价格优势。

259　信息栏：细菌微阵列

在大约 50 年的过程中，细菌遗传学家已经对 *E. coli* 中几乎全部的基因和操纵子进行了作图、测序、突变和移位，并将它们与表型联系起来。他们对包含特定代谢功能的基因簇进行了分类，并且鉴定了生理变化的传感器，发现了用于开放和关闭相关基因库的应答调控途径。如果所有的这些已经为权威的遗传学和生物化学所实现，那么还需要用 DNA 微阵列去研究原核生物中的基因表达吗？

当研究人员可以应用选择压力来增强或者是抑制所感兴趣的基因功能时，遗传学可以发挥其最大的用处。未与表型相联系的、带有未知功能的基因通常被遗传学家忽略或者视而不见。这在 1997 年变得实在太明显，当 *E. coli* 基因组全部 4.6 Mb 序列公布时（Blattner et al. 1997），*E. coli* 染色体中不少于 30%的可读框被发现带有未知的功能。当 *E. coli* 在不同的生理环境和遗传状态之下生长时，DNA 微阵列这种可以定量每个基因的 RNA 水平变化的能力，为分析基因组对于特定生理学反应的全局性响应和新的基因指派功能提供了可能性。

首次针对 *E. coli* 基因表达进行全局性分析的尝试发表于 1993 年，实验使用了一个有序的、细菌噬菌体 λ 基因组重组体的重叠集，按照 21×21 的矩阵排列在尼龙膜上（Chuang et al. 1993）。由于重组体克隆中的插入片段很大（9～21 kb），包含了数个基因，按照现在的标准系统分析的分辨率是低的。然而，它可能鉴别包含了表达水平受到多种环境变化，包括热休克、渗透休克、各种营养匮乏、低氧张力和化学诱导剂等影响的基因的区域。虽然在全部的基因组序列得到后，更详细的表达谱分析是可能的，但原

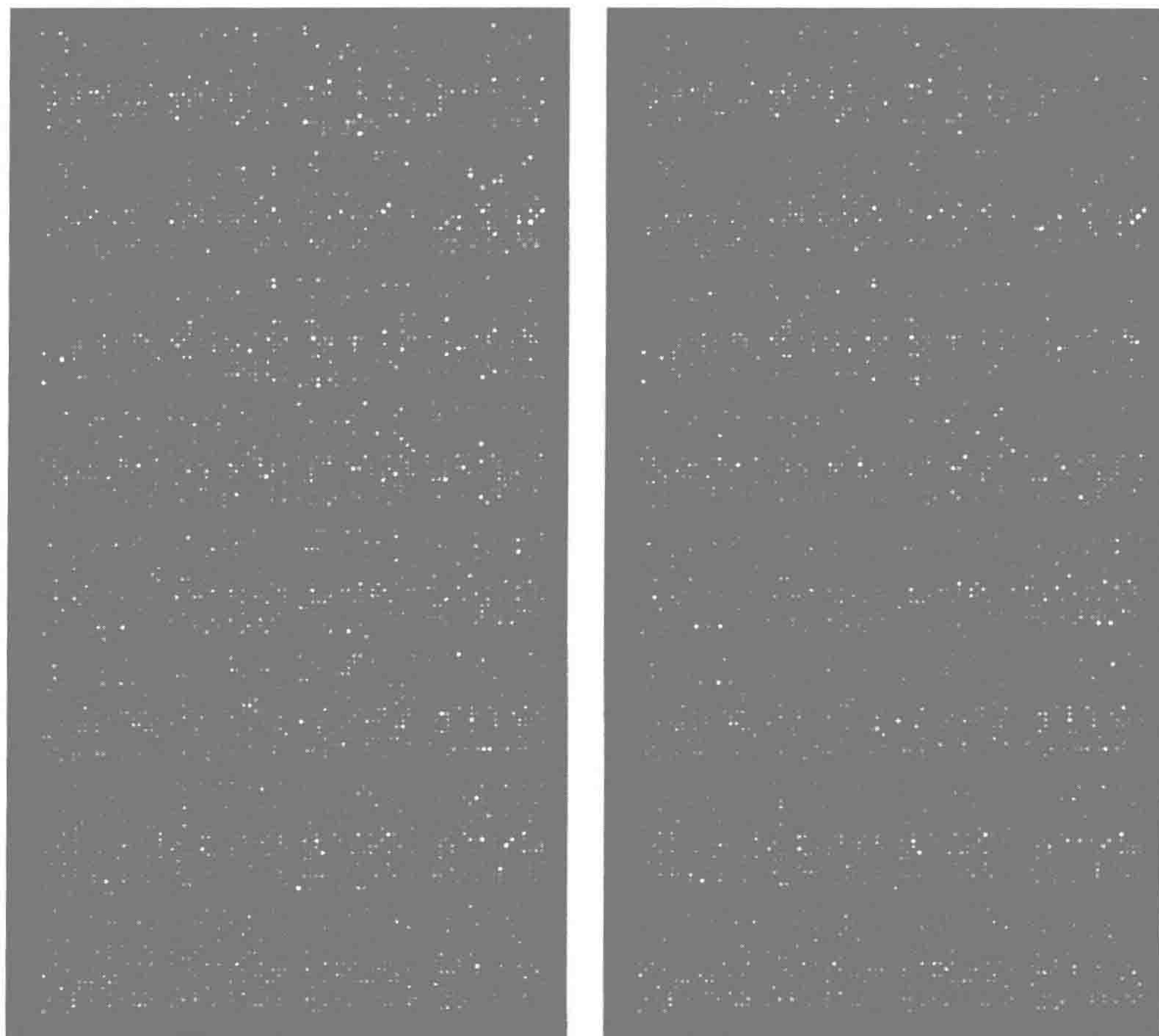


图 3-25 Alex-Fluor 546-14-dUTP 和 Cy-dUTP 染料的比较

来自野生型 *Arabidopsis thaliana* 的 RNA，用 Alex-Fluor 546-14-dUTP(Molecular Probes)（左图）或者 Cy3-dUTP(Amersham)（右图）标记，与 Cy5 标记的（Amersham)*A. thaliana* cim-7 突变体竞争杂交。每种染料的掺入实验使用相同的 RNA，在一组三个微阵列上进行。两类玻片均使用标准的第二链 cDNA 标记反应（方案 13 和 http://afgc. stanford. edu/afgc _ html/AFGCProtocols-fev2001. pdf）和杂交方法（方案 20）。使用 Packard Bioscience Scanarray 4000(Perkin Elmer) 进行扫描。Cy3-dUTP 和 Alex-Fluor 546-14-dUTP 通道的图像扫描时激光功率为 77，PMT 为 88，而采集 Cy5 通道时激光功率为 70，并且 PMT 为 70。染料给出了近似结果（图像由 Lorne Rose 提供，Carnegie Institute of Washington）（见彩图）

核生物的全局性基因表达分析滞后于真核生物中的相近似的研究——也许是由于细菌 mRNA 短暂的半寿期和缺少通过 poly（A）尾巴进行的分离，或者是存在压倒多数的 rRNA 和 tRNA 的情形下选择性的标记 RNA。这些问题大体上已经得到解决，并且在最近几年中，关于原核生物表达分析的论文数量在稳步增长。

用于原核生物表达分析的大多数实验与布朗实验室中为酵母而研究的实验，在设计和实施上大体相似。阵列由基因组元件的详细集组成，实验设计上主要包括比较对照 RNA 和一系列分离自暴露于药物、诱导剂或者各种形式的代谢和环境胁迫下的培养物中的实验组 RNA 各自的靶序列丰度。数据的收集与分析和酵母表达谱实验收集的数据

处理方式相同。

理论上，DNA 微阵列可用来对基因组已经被完全测序的许多原核生物进行全局性基因表达分析。然而，当有关生物的丰富遗传知识可以得到，并且突变体的表型得到很好的鉴定时，对微阵列数据的解释不但被简化，而且更加丰富。在本书写作的时候，有大于 20 种原核生物的全部基因组序列已知，但关于表达分析的绝大多数工作仍然集中在 *E. coli* 上，如同下面所总结的那样。

260

- **阵列的构建**。最常用的阵列形式由对应于每个可读框的 PCR 产物组成，点样到载玻片上。以 *E. coli* 为例，使用特异的引物对，从基因组 DNA 模板（30 ng）中扩增了 4290 个可读框（Richmond et al. 1999）。引物的设计通常以公认的起始密码子 ATG 开始，并且在终止密码子处结束。如果需要的话，引物可以设计为 5′序列包含一个限制性的位点以便于克隆。扩增通常分两步进行，以减少点样到微阵列上的模板 DNA 的量。第一次扩增反应的产物稀释约 500 倍，并用作再次扩增的模板。在使用 96 孔 PCR 纯化试剂盒（QIAGEN）进行纯化之前，第二轮 PCR 产物的长度和产量通过凝胶电泳来确认。一些实验中（Richmond et al. 1999），99.5%的 *E. coli* 可读框能够成功地扩增并用于微阵列实验（Wei et al. 2001）。阵列上的基因可以被分类成 Riley 和 Labedan 定义的不同功能集合（1996；Smulski et al. 2001；Wei et al. 2001）。

  Selinger 等（2000）介绍了一种不同的方法，将 *E. coli* 基因组的编码区和非编码区呈现在由具有特定序列的 25 聚体寡核苷酸构成的 544×544 方格中。整个阵列包含了 295 936 种寡核苷酸，在衍生的玻片表面原位合成（Pease et al. 1994；Lockhart et al. 1996）。数量巨大的寡核苷酸能够用来高分辨率的检测转录物（平均基因间隔区为每 6 个碱基，可读框为每 60 个碱基）。尽管该系统的灵敏度几乎与通常的点样微阵列相当（每个细胞约 0.2RNA 拷贝），但它的高分辨率和能够检测非编码区转录物改变的特点对一些研究人员来讲是有价值的。

- **标记 RNA**。通常使用 QIAGEN RNeasy 分离柱从 *E. coli* 制备用于表达分析的 RNA（Richmond et al. 1999；Khodursky et al. 2000），但一些研究组的人员宁可使用更加精细的方法（Wei et al. 2001）。总 RNA（6～50 μg）可以应用商品化的随机引物、反转录酶和 Cy3-或 Cy5-标记的脱氧核苷酸，在随机延伸反应中进行标记。作为替代，在由 poly(A) 聚合酶催化的体外反应中，细胞总 RNA 存在的情形下，细菌 mRNA 的 3′端可以被选择性地聚腺苷酸化修饰。聚腺苷酸化的分子可以使用 oligo(dT) 色谱柱进行纯化（Wendish et al. 2001）。标记和杂交后，从总 RNA 和约 1%的经过 oligo(dT) 纯化的 mRNA 中得到的信号强度相当。从两种类型的靶分子得到的信号比值几乎是相等的，这表明 mRNA 的富集可能没有显著的优势。

  可以选择的标记方法包括末端转移酶催化的第一链 cDNA 的生物素化修饰（应用于与 Affymetrix 类型阵列的杂交；Selinger et al. 2000）和用于选择性标记转录物的针对基因组的引物（Talaat et al. 2000）。这种方法使用计算机算法来确定最少数量的、具有固定长度且能够从全部已知基因的基因组拷贝启动合

成的寡核苷酸。

- **数据的标准化**。从荧光标记 cDNA 获得的杂交信号可除以来自荧光标记的切断的基因组 DNA 所产生的信号而被标准化（Richmond et al. 1999；Wei et al. 2001）。此外，大多数研究人员在他们的微阵列中包含了一组用作内参比的 DNA。例如，Wei 等（2001）在每张载玻片上点样 76 种特异性的 *E. coli* PCR 产物，8 种扩增的 *Klebsiella pneumoniae* 的基因和 12 种植物 cDNA 克隆。内参比用作在构成整个基因组的阵列载玻片之间获得相等的读数。 261

表达谱分析是迄今为止使用 DNA 微阵列研究原核生物的最常用的方法，阵列技术也应用于比较研究，对紧密相连的原核生物基因组之间的基因特异性差异进行检测和快速作图，并研究显示它们演化关系的系统发生树（Murray et al. 2001）。

信息栏：用酚抽提蛋白质和 DNA

直到 20 世纪 50 年代中期，纯化 DNA 的标准方法包含了使用去垢剂和强烈的盐溶液（如高氯酸盐）从核酸上剥离蛋白质。最终的蛋白质清除通过使用添加有少量异戊醇的氯仿进行数次抽提而实现（Sevag et al. 1938）。第一次报道的使用酚来纯化核酸的方法由 Kirby(1956) 发表，他觉察到酚从水溶液中抽提蛋白质的能力（Grassmann and Defner 1953）。在他最初的文章中，Kirby 展示了使用两相的酚-水混合物抽提哺乳动物组织匀浆，使得在室温下 RNA 分配进入水层。在界面上，DNA 仍然和蛋白质结合在一起。Kirby 迅速意识到用阴离子盐溶液替代水可以将 RNA 和 DNA 都释放到水相中（Kirby 1957；Kirby 1964）。尽管使用阴离子盐从 DNA 上释放蛋白质的方法很快被抛弃，并代之以强的阴离子去垢剂，如 SDS，Kirby 关于酚抽提的最初描述构成了许多目前广为使用的纯化方法的基础。酚的功能可能是作为蛋白质的溶剂，因为它可以抽提被阴离子盐或者去污剂从核酸中释放的蛋白质。该过程非常高效，以至于只需要用酚抽提两、三次就可以得到核酸的纯的制备物。

酚的 pH 决定了 DNA 和 RNA 在有机相和水相之间的分配（Perry and Kelley 1972；Brawerman et al. 1972）。在弱碱性的 pH 下，DNA 和 RNA 都几乎定量地分配到水相中；在酸性的 pH（如 pH 为 5.2）下，DNA 发生变性并沉淀到有机相和两相的界面中。另一方面，RNA 仍然保留在水相之中。因此，分离没有显著 DNA 污染的 RNA 需要酚或者酚：氯仿的酸性溶液。在使用之前应检查生产商提供的酚的 pH，并用 0.01 mol/L Tris(pH＜7.00) 平衡至适当的值，如果是酸性酚，在 10 mmol/L 醋酸钠（pH 5.1）、50 mmol/L NaCl 和 1 mmol/L EDTA(pH 为 8.9) 中平衡。盛放平衡酚的瓶子应该用铝箔包裹，并在 4℃下储存。

经过纯化的酚的比重为 1.07，因此当它和水混合的时候形成下层相。然而，使用酚从含有高浓度溶质的溶液中抽提蛋白质时，有机相和水相可能难以分离或者倒置。使用 50：50 的酚：氯仿混合物时，这一问题可以在很大程度上得到缓解，因为氯仿的高密度（1.47）确保了两相的分离。变性的蛋白质集中在两相界面处，而脂质被有效地分配到有机相中。常常向酚：氯仿混合物中加入异戊醇以减少泡沫的产生。有机溶剂的联合使用被认为可以减少 RNA 向有机相的分配和在界面处形成不溶的 RNA-蛋白质复合物。 262

纯品酚为白色晶体（熔点 43℃）。暴露于空气和光后，会逐渐变红，这一过程在碱性条件下会加快。带有颜色的晶体酚批次必须在 182℃下重蒸以除去会导致磷酸二酯键断裂或促进核酸交联的氧化产物，如醌类物质。

许多供应商提供的液化的酚含有约 8%的水，可以冰冻保存在－20℃。液化的酚如果是无色的，可以不用重蒸而直接应用于分子克隆。现在，偶尔会有一些批次的液化酚带有粉红色或者黄色，这样的酚应该拒绝接受并退还生产商（饱和酚可以从许多生产商得到，包括 Ambion）。

检查平衡酚的 pH

- 酸性酚：将 2 ml 有机相、8 ml 乙醇和 10 ml H_2O 混合，用标准的参比电极测量最终溶液的 pH。
- 中性酚：将 2 ml 有机相、5 ml 乙醇和 13 ml H_2O 混合，用标准的参比电极测量最终溶液的 pH。

263

信息栏：甲酰胺

甲酰胺被用作水溶液缓冲液的离子化溶剂。大多批次的高品质甲酰胺纯度已经足够，可以直接使用而不需要进一步处理。然而，由经验得知，如果出现黄色或者轻微的氨的味道，应当纯化甲酰胺。一个更严格的检查纯度的方法是测量电导率，它随着甲酰胺分解为甲酸铵而上升。纯净的甲酰胺的电导率为 1.7(Casey and Davidson 1977)，浓度为 10^{-3}mol/L 的甲酸铵的电导率为约 650 $\mu\Omega$。用于复性实验中的甲酰胺的电导率应该小于 2.0 $\mu\Omega$。

甲酰胺和混合床离子交换树脂［如 Dowex AG8，Bio-Rad AG 501-X8，20～50 目或 X8(D)］一起放在磁力搅拌器上搅拌 1 h 可以去离子化。之后将溶液用 Whatman ＃1 滤纸过滤并分装成小份保存在－20℃下，最好处在氮气氛围下。每种树脂可以重复使用数次。X8(D) 带有指示剂，当树脂耗尽时颜色发生改变。

如下所述，杂交反应中使用甲酰胺来解决测序胶中复杂的压缩，并在电泳前使 RNA 变性。

电泳前使 RNA 变性

甲酰胺（50%）用于在 RNA 通过变性甲醛-琼脂糖凝胶电泳之前帮助 RNA 的变性(Lehrach et al. 1977)。

杂交反应

Bonner 等（1967）首先使用甲酰胺作为杂交反应的溶剂，在他们简短的文章结尾之处写到：

> 甲酰胺取代杂交过程中升高的温度是能够期望的。正如 Helmkamp 和 Ts'o(1961) 以及 Marmur 和 Ts'o(1961) 所显示的，甲酰胺的水溶液使 DNA 变性。DNA-RNA 杂交体需要的甲酰胺浓度为 30%～40%（体积比），比 Marmur 和Ts'o发现的使自

然 DNA 变性时（在 0.02 mol/L NaCl-0.002 mol/L 柠檬酸钠中）的 60%（V/V）要低很多。

现在已经幸运地发现，在甲酰胺水溶液中进行的杂交反应与在升高的温度下进行的杂交反应相比具有显著的优点。这些优点包括：提高了硝酸纤维素膜对固定的 DNA 的保持能力，降低了非特异性背景吸附。这两个因素相结合，导致杂交过程重复试验的重现性增高。低温条件下在甲酰胺溶液中进行的杂交，也有助于减少长时间孵育过程中核酸分子的断裂。

除了这些优点以外，还为给定实验反应条件的设计增加了灵活性。使用甲酰胺使得杂交严格度的控制更加方便，而不需要借助于调节杂交的温度来实现。

264

在含有甲酰胺的缓冲液中进行杂交

双链 DNA 解链温度（T_m）的下降是甲酰胺浓度的线性函数（McConaughy et al. 1969；Casey and Davidson 1977）。DNA 中的 G+C 含量为 30%～75%时，杂交混合物中每百分之一的甲酰胺使得 T_m 下降 0.63℃。因此，探针和靶分子形成的杂交体的 T_m 可以用下面的公式来估计，该公式根据 Bolton 和 McCarthy(1962) 并进行了修改：

$$T_m = 81.5℃ + 16.6(\lg[Na^+]) + 41(\text{摩尔分数}[G+C]) - 0.63(\text{甲酰胺}\%) - 500/n$$

式中，n 是用核苷酸表示的 DNA 的长度。该公式适用于以光学方法测量在 OD_{260} 处增色现象而定义的可逆的 T_m。对于使用放射自显影的方法检测 DNA 杂交体更为重要的“不可逆”T_m（Hamaguchi and Geiduschek 1962），通常要比由公式预测的高7～10℃。RNA 探针与固定化的 RNA 杂交也有类似的公式（Bodkin and Knudson 1985）：

$$T_m = 79.8℃ + 18.5(\lg[Na^+]) + 58.4(\text{摩尔分数}[G+C]) + 11.8(\text{摩尔分数}[G+C])^2 - 0.35(\text{甲酰胺}\%) - 820/n$$

而 DNA-RNA 杂交体为（Casey 和 Davidson，1977）：

$$T_m = 79.8℃ + 18.5(\lg[Na^+]) + 58.4(\text{摩尔分数}[G+C]) + 11.8(\text{摩尔分数}[G+C])2 - 0.50(\text{甲酰胺}\%) - 820/n$$

比较这些公式表明，在高浓度甲酰胺条件下，核酸杂交体的相对稳定性按照如下的顺序递减：RNA-RNA（最稳定）、RNA-DNA（稳定性稍差）、DNA-DNA（最不稳定）。在 80%的甲酰胺中，RNA-DNA 杂交体的 T_m 要比具有相同碱基组成的DNA-DNA杂交体高出约 10℃。因此，能够找到一个杂交条件使得可以形成RNA-DNA杂交体，而不能形成 DNA-DNA 杂交体（Casey and Davidson 1977）。这种可以抑制 DNA 复性的性能在用双链 DNA 探针进行 RNA 的 S1 作图时非常有用（Berk and Sharp 1977）。然而，现在已经研究出制备单链探针的高效方法，使得 RNA 和 DNA 的复性可以在标准的杂交条件下进行，而不用担心来自互补的 DNA 链的竞争。因此，由 Casey 和 Davidson(1977) 建立的复性条件现在几乎不再使用。

- 在 80%的甲酰胺中，DNA-DNA 杂交的速度要比在水溶液中慢（Casey and Davidson 1977）。升高甲酰胺的浓度会降低 DNA：DNA 的复性速度，甲酰胺的浓度每升高 1%，杂交速度下降 1.1%（Hutton 1977）。因此，在 50%的甲酰胺中最理想的速度是水溶液中最理想速度的 0.45 倍。在 80%甲酰胺中，DNA-DNA 杂交的速度要比水溶液中慢3～4 倍（Casey and Davidson 1977），这种影响是在用于复性的温度下杂交溶液黏度增高的结果。
- 在长时间温度超过 37℃的孵育下，甲酰胺发生分解，导致杂交缓冲液的 pH 增高（Casey and

Davidson 1977)。杂交缓冲液中包含了甲酰胺的时候，6×SSPE 比 6×SSC 更好，因为它具有较强的缓冲能力。

（翻印自 Sambrook and Russell 2001）

265

信息栏：采用胍盐抽提 RNA

从动物和植物细胞中纯化 RNA 的最新方法是基于胍盐，它是最有效的蛋白质变性剂和核糖核酸酶抑制剂之一。在分离 RNA 的实验中，最初用作脱蛋白试剂的胍盐是盐酸盐（Cox 1968）。在 20 世纪 70 年代，一系列的文章介绍了使用更强的异硫氰酸盐从培养细胞和新鲜组织中分离总 RNA(Glisin et al. 1974；Ullrich et al. 1977；Chirgwin et al. 1979)。80 年代早期，这些方法取代了陈旧的、基于有机溶剂的抽提技术，从“难以处理的”来源，如富含核糖核酸酶的胰腺组织中纯化 RNA。然而，早期基于胍盐的方法只能同时处理少数样品，因为它们使用了细胞和组织裂解物的 CsCl 密度梯度超离心来分离残留的细胞组分中的 RNA。直到 1987 年，Chomczynski 和 Sacchi 克服了对超速离心的需求，基于胍盐的方法可以同时处理比超速离心机转子所能容纳的更多的样品。

进一步的修改包括应用 0.5 倍体积的异丙醇和 0.5 倍体积的高盐溶液从细胞裂解物中回收 RNA，这一改进消除了最终 RNA 制备物中的多糖污染物（Chomczynski and Mackey 1995)。根据这种方法研究的最初的商品化试剂——RNAzol 和 RNAzol B(Cinna Scientific，Cincinnati，Ohio)，提供了一个从多种生物样品中获得高质量的、未降解 RNA 的快速、价廉和简便的方法。在 20 世纪 90 年代发表的一系列文章中，Chomczynski(1993，1994，1999) 介绍了一种能够同时回收 RNA、DNA 和蛋白质的修改方法。在这个方法里，细胞和组织用异硫氰酸胍、硫氰酸铵和酚的单相溶液，以及 TRI 或者 TRIzol 试剂进行裂解，这些试剂，现在可以从不同的公司得到，商品名称各异（见本部分 PART Ⅰ 中表 3-1 的脚注 c)。向细胞裂解物中加入氯仿产生第二（有机）相，DNA 和蛋白质被抽提到其中，而 RNA 留在上层的水中。如果仔细操作避免产生界面层，从水相中用异丙醇沉淀回收的 RNA 可以不做进一步的纯化而直接与 DNA 微阵列杂交或者作为完整的 poly(A)$^+$RNA 的来源。然而，一些研究人员更愿意使用如 FastTrack 2.0 试剂盒（Invitrogen）来分离 poly(A)$^+$RNA（见方案 2)。

含有异硫氰酸胍的单相裂解试剂能够非常高效地裂解细胞和变性蛋白质，而且它们也是从全部动物细胞和大多数植物细胞中分离 RNA 的首选缓冲液。然而，对于一些特定的植物组织，使用含有盐酸胍的裂解试剂效果更佳。例如，在异硫氰酸胍存在的条件下，玉米的胚乳产生严重的沉淀，但在含有盐酸盐的裂解液中却是可溶的。QIAGEN 出售的 RNeasy Plant Mini 试剂盒含有两种裂解溶液，并给出了从不同的植物组织之中抽提 RNA 时采用何种裂解液的建议。

生产商提供了许多种单相裂解试剂的试剂盒（见本部分 PART Ⅰ 中表 3-1 的脚注 c)。一些单相裂解试剂的详细组分在公开的专利中有介绍（Chomczinski 1989，1994)。

信息栏：动物细胞及组织的匀浆

266

下面是两种类型常用的电动匀浆器。

- **The Kinematica Polytron** PT3100 型仪器装备有 9100071 探针（Brinkmann），用于破碎大于 100 mg 的细胞和组织。尽管 Polytron 每次只能处理一份样品，但它配备的强力且可调速的马达使得一次操作中可以对数克重的样品进行匀浆。匀浆前样品不需要解冻。Polytron 的缺点是在样品之间可能会发生交叉污染。由于基因表达水平的差异有几个数量级，即使是先前样品裂解物对匀浆物的轻微污染，也会影响到 DNA 微阵列上的基因表达谱。要在不同的样品之间清洁 Polytron：
    1. 拆开探针，除去在转头和刀片部分全部的残留组织。将转头和刀片在大量的水中煮沸数分钟。水在探针的内部会沸腾，因此将整个探针没入大量的水中是非常重要的，且水要更换一两次。
    2. 在 50 ml 的新鲜水中运行重新装配好的 Polytron。
    3. 用干净的 Kimwipes 纸巾擦干 Polytron 探针。
- **The FastPrep instrument**（Q-Biogene，6000-120）用于破碎小于 100 mg 的组织，并能够同时处理最多 12 个样品。由于每个样品放在各自的独立试管中，减少了样品交叉污染的机会。根据不同的组织类型，可能需要运行多次以获得完全的匀浆。当样品的脂肪含量高时，需要大量的样品以获得足够数量的 RNA。关于更多的细节，见方案 2。

虽然电动匀浆器大体上取代了组织破碎的手工方法，但下面列举的老式的技术仍然适用于从骨骼或者软骨组织、植物材料、富含降解酶的组织中制备 RNA。

当处理如胰腺或者是内脏这些富含核糖核酸酶的组织时，速度是成功的关键所在。最好将新鲜切除的组织切碎成小块（约 100 mg），并将它们迅速浸入液氮中，快速冰冻的组织可以在－80℃下保存数个月而不会影响到 RNA 的产量或者完整性。要抽提 RNA：

1. 将约 100 mg 冰冻的组织转移到在液氮中预冷过的研钵，并用预冷的研杵将组织研磨成粉。在研磨过程中加入液氮可以使组织保持冷冻状态。使用较大的研钵，经常向上面加入液氮，防止加入液氮时粉末飞溅出来。
2. 将组织粉末转移到 TRIzol 溶液（1 ml/100 mg 组织粉末）中，在 Polytron 匀浆器中处理（方案 1）或者按照 100 mg 的批次在 FastPrep 仪器中处理（方案 2）。

267

速冻和研磨并不总是必要的。核糖核酸酶含量不是很高的组织可以切成小块，然后立即在 TRIzol 中裂解或者在 RNA*later*(Ambion) 中短期保存（见表 3-1）。

非常硬的组织，如骨骼和软骨应该使用经过液氮预冷的研钵和研杵研成细粉末，然后将粉末在适量的 TRIzol 溶液中用 Polytron 匀浆器匀浆 30 s。在这个过程中，组织完

全溶解于 TRIzol，而且 DNA 被切碎。匀浆结束后，将 Polytron 探针在装有 3 ml TRIzol溶液的 15 ml 锥形试管中运行，然后将 TRIzol 合并到组织裂解液当中。如果仍然有可见的颗粒状物质，将样品在 Sorvall 离心机（或者与之相当的设备）中，于 4℃以 12 000 *g* 离心 30 min。

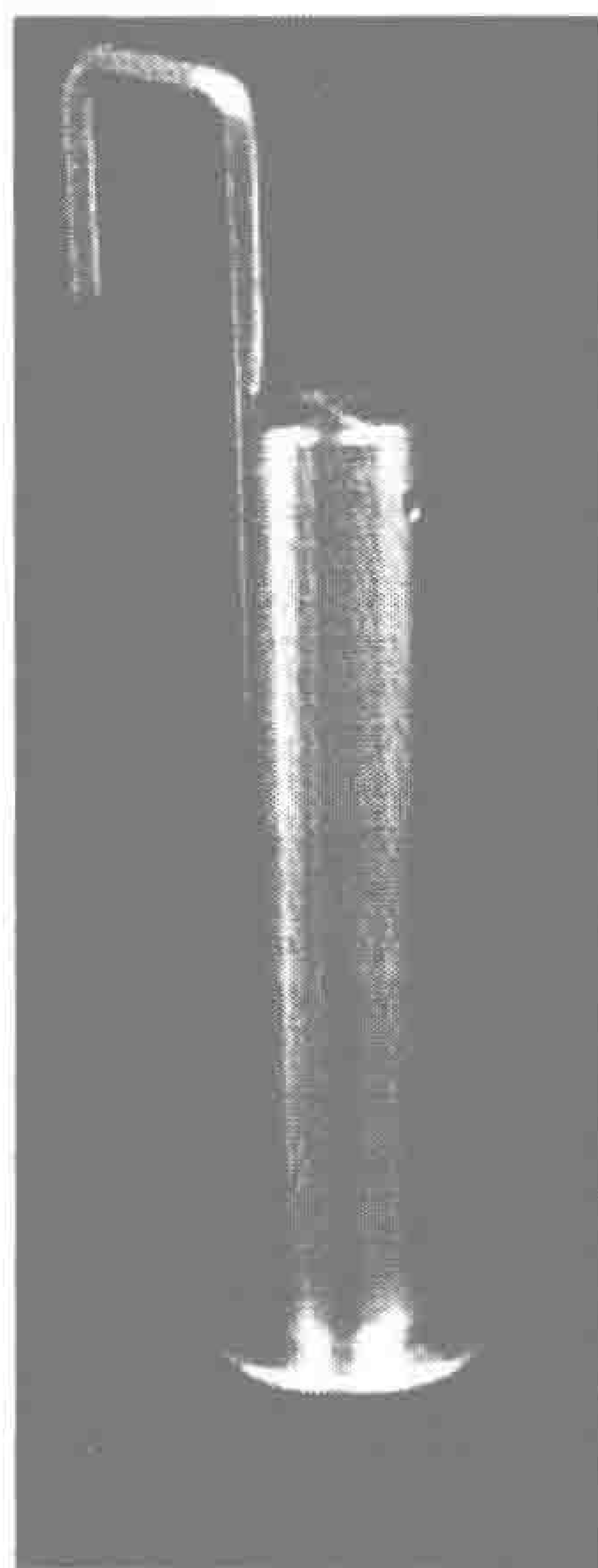


图 3-26　不锈钢活塞式匀浆器

冰冻组织除了在研钵中磨碎外，还可以放入自制的塑料薄膜袋中，并用钝的设备，例如锤子研成粉末（Gramza et al. 1995）。只有特定类型的塑料薄膜才有足够的强度，能够承受住低温下的锤打，例如，3M 的 Write-On Transparency Film AF4300。更进一步的修改是使用在干冰上预冷的不锈钢制的活塞和研钵（见图 3-26）。拥有许多这些自制的仪器可以加快多组样品的处理。

这里讨论的方法是为处理动物细胞和组织而研究的。匀浆植物组织需要不同的技术（见方案 2 和方案 3）。

信息栏：抑制 RNase

RNase 是稳定性高且强有力的酶，在 RNA 分离和表征的所有阶段都会严重威胁到 RNA 的完整性。下面介绍了三种常用的抑制剂，用于阻止 RNase 的活性。

焦碳酸二乙酯

DEPC 是高反应活性的烷化试剂，用于缓冲溶液和玻璃器皿中 RNase 活性的抑制（Penman et al. 1970；Williamson et al. 1971）。由于 DEPC 不加选择地修饰蛋白质和 RNA，它在分离和纯化 RNA 时不能使用，并且与一些类型的缓冲液不相容。DEPC 是具有很高反应活性的烷化试剂，主要通过对组氨酸残基的乙氧基甲酰化（ethoxyformylation）来破坏 RNase 的活性。

在水溶液中，DEPC 迅速水解为 CO_2 和乙醇，在 pH 6.0 的磷酸盐缓冲液中半寿期约为 20 min，而 pH 7.0 时为 10 min。这个水解过程可以被 Tris 和其他胺加速，而它们自身在这个过程中被消耗掉，因此 DEPC 不能用于处理包含这些缓冲液的溶液。不含有亲核试剂（如 H_2O 和乙醇）的 DEPC 样品非常稳定，但是很少量的这些溶剂就会使其完全转化为二乙基碳酸盐。因此，DEPC 保存时应该防潮。将 DEPC 分装成小份并在干燥条件下保存，并且永远不要在试剂瓶的温度低于环境温度时打开瓶子。

要使玻璃器皿去除污染，将器具充满 0.1%的 DEPC 水溶液，并在 37℃下放置2 h。用不含 RNase 的水清洗器具数次，然后以高压蒸汽灭菌方式（liquid cycle）将它们在 15 psi(1.05 kg/cm^2) 下灭菌 15 min。

通过维护良好的、现代的反渗透系统纯化的 H_2O 是不含 RNase 的（Huang et

al. 1995)。然而，维护情况不好的纯化系统可能受到微生物的污染。这在具有数米长的管道系统和储液桶的大型中央系统中是常见的问题，水常会在其中滞留。在这类情况下，按照上述的方法，用 DEPC 处理来得到不含 RNase 的水是必需的。

DEPC 可使 RNA 中未配对的嘌呤残基羧甲基化。在体外蛋白质合成体系中，暴露于 DEPC 的 mRNA 的转录活性下降（Ehrenberg et al. 1976）。然而，DEPC 处理过的 RNA 形成 DNA-RNA 或者 RNA-RNA 杂交体的能力受此影响却不大，除非很大一部分的嘌呤残基被修饰。

加热降解除去 DEPC 会产生少量的乙醇和 CO_2，这会使得未缓冲的溶液离子强度增高且 pH 降低。不要在含有游离氨基的溶液（如 Tris 缓冲液）中使用 DEPC。

注意事项：DEPC 具有毒性并可能致癌。穿戴手套并佩戴防护眼镜，防止 DEPC 接触到皮肤和眼睛。在通风良好的化学通风橱中操作该化学试剂。

该信息栏根据 Sambrook 和 Russell(2001) 修改而来。

氧钒核糖核苷复合物

这些复合物是过渡态的类似物，可以与许多 RNase 的活性部位结合而几乎完全地抑制它们的催化活性（Berger and Birkenmeier 1979）。由于氧钒核糖核苷不是共价修饰 RNase，它们必须在 RNA 抽提和纯化的全部阶段一直使用。然而，由于这些复合物抑制 RNA 聚合酶和体外转录，它们必须通过使用含有 0.1% 羟基喹啉的酚进行多次抽提而被从最终的 RNA 制备物中除去。多家供应商提供氧钒核糖核苷复合物。

RNase 的蛋白质抑制剂

许多 RNase 与一种在几乎全部哺乳动物组织的细胞质中都能发现的大约 50kDa 的蛋白质产生非常紧密的结合，虽然这种结合方式不是共价的。从胎盘中可以分离到大量的这种蛋白质（Blackburn et al. 1977）。在体内，这些蛋白质抑制剂隶属于胰腺的 RNase 超家族，特别值得注意的有血管生成素（angiogenin），它是一种血管诱导因子和叶曙红素衍生的神经毒素。这些蛋白质抑制剂和它们的靶分子之间的 1∶1 的复合物，是有记录的、结合最紧密的复合物之一（1～70 fM；Lee et al. 1989；Lee and Vallee 1993）。

从多种来源获取的 RNase 的蛋白质抑制剂由数家生产商销售，商品名称各不相同（如来自 Promega 公司的 RNasin 和 5 Prime→3 Prime 的 Prime Inhibitor）。虽然它们对巯基试剂的需求各不相同，但却显示了对 RNase 的广谱抑制活性，且不会抑制其他核酸酶、聚合酶或者体外翻译体系（Murphy et al. 1995）。

因为这些抑制剂不与 RNase 之间形成共价复合物，所以他们不能用于存在变性试剂如 SDS 和胍的场合，这些变性剂常用于在 RNA 抽提的最初阶段裂解哺乳动物细胞。然而，这些抑制剂可用于随后的 RNA 纯化的全部过程中。在纯化过程中，抑制剂必须补充数次，因为它们会被酚抽提除去。

270 信息栏：哺乳动物、植物和细菌的 RNA

哺乳动物 RNA

一个典型的哺乳动物细胞包含约 $0.5\times10^{-5}\sim1.05\times10^{-5}$ μg 的 RNA，其中大部分在细胞质中（表 3-10）。四种 rRNA（5S、5.8S、18S 和 28S）占到细胞质 RNA 的 80%～85%，而全部的 tRNA 占到大约 15%。高丰度的 RNA 具有固定的大小和序列，并能够通过凝胶电泳、HPLC、阴离子交换色谱或者密度梯度离心分离成几乎纯净的形式。与此形成对比的是，mRNA 只占到总的细胞 RNA 的 1%～5%，分子大小不等（长度从数百个碱基到数千个碱基），丰度各不相同（从每个细胞中 1 或 2 个分子到多达数千拷贝），序列差异也很大。大多数哺乳动物的 mRNA 的 3′端携带有聚腺苷酸残基区域，通常足够地长，因而可以使用 oligo(dT)-纤维素上的亲和层析来纯化 mRNA [见 **oligo(dT)-纤维素**的信息栏]。由此得到的各种 mRNA 分子的集合共同编码了细胞内几乎全部的多肽。

表 3-10 哺乳动物细胞质中 RNA 的组成

RNA 的类型	长度/(核苷酸数)	近似分子质量/Da[a]	细胞中的种类	占细胞质 RNA 的重量百分比/%
18S 核糖体	1868[b]	0.58×10^6	1	约 20
28S 核糖体	5025[b]	1.56×10^6	1	约 57
5.8S 核糖体	158	4.90×10^4	1	约 1.5
5.0S 核糖体	120	3.70×10^4	1	约 1.5
tRNA	73～91	不均匀的	约 100	15～17
mRNA	数百到数千	不均匀的	$>10^4$	1～5
小 RNA	100～330	不均匀的	约 60	<2

a 使用每个核苷酸的相对分子质量为 310 进行计算。

b 给出的长度为人的 18S 和 28S 核糖体 RNA。

植物 RNA

植物 RNA 由非均一的分子构成，可以通过区带沉降分为三类：23S～25S、16S～17S 和 4S。前面两类由 rRNA 所组成，而 4S 类则包含了 tRNA 和其他小的 RNA 分子，如小核 RNA(Brown and Shaw 1998；Sugiura and Takeda 2000)。这三类 RNA 丰度高而且稳定，并且在经过溴化乙啶染色的琼脂凝胶上很容易观察到，呈现为具有很强荧光的区带。rRNA 是在全部光合真核生物，包括植物和藻类中可以发现的三种核糖体（即细胞质、质体和线粒体）的组成成分。

271 mRNA 在植物细胞中的丰度远不如在动物中的，通常只占到细胞总 RNA 的 1%～2%。和动物的 mRNA 一样，植物的 mRNA 在序列、稳定度和长度上也是不均匀的，从数百个碱基到数千个碱基（Spremuli 2000）。一个典型的植物细胞含有稍少于 1 pg 的 mRNA，并取决于细胞的类型、发育阶段和生长的条件（Okamura and Goldberg 1989）。假设 mRNA 的平均长度为 2 kb 并且一个细胞中有 0.6 pg 的 RNA，那么一个植

物细胞可能含有约 5×10^5 个 mRNA 分子。此外，mRNA 的表达水平有较大变化，从高丰度的，如特定植物组织中的光合作用基因，到中等水平和稀少的。例如，编码核酮糖-1，5-二磷酸羧化酶/加氧酶（rubisco）全酶小亚基的转录物在番茄中占全部转录物的大约 3%（Pichersky et al. 1986）；编码亲环素（cyclophilin）的约占 0.3%；编码肌动蛋白的，在番茄叶中约占 0.04%(Karrer et al. 1995)。mRNA 在细胞核中产生，其中一些在进入细胞质之前要经过转录后处理。由于植物细胞具有坚硬的细胞壁，大多数分离 RNA 的方法需要将整个细胞粉碎，因此得到的 RNA 合并物中可能包含了不均匀的细胞核内的转录物（即未经处理的和处理了一半的转录物）。为了只得到细胞质的 RNA，考虑从温和裂解的植物原生质体中抽提，该方法使得完整的细胞核可以被分离出来并丢弃。

植物 mRNA 3′端的 poly(A)$^+$ 尾巴使得它们明显地区别于其余的各种 RNA 分子，并且这一独特的性能使得我们能够将其纯化并与其他种类的 RNA 分离，正如在其他真核生物中的那样。

细菌 RNA

E. coli 细胞中大约 4%的总 RNA 是 mRNA，而大约 80%的为 rRNA(16S 和 23S)，15%的是 tRNA。*E. coli* 基因组的序列预测有 4290 个可读框，其中约 70%的已经被指派了功能（Blattner et al. 1997）。与真核生物的 mRNA 不同，原核生物 mRNA 没有全部聚腺苷酸化，因此不能使用 oligo(dT) -纤维素层析进行纯化。细菌 mRNA 短暂的半寿期，无法在具有更高丰度的 tRNA 和 rRNA 存在的情况下对他们进行选择性的标记，这无疑导致了微阵列技术作为一种研究全局性基因表达的方法在原核生物中的渗透相对迟缓。然而，这些问题并非不能克服，而且细菌表达谱的脚步正在迅速地加快（关于进一步的详细情况，见**细菌微阵列**信息栏）。

质量控制

不论来源如何，用于表达分析的 RNA 制备物的质量应该使用凝胶电泳进行检查(图 3-27 和图 3-28)。检查哺乳动物和植物 RNA 完整性的其他方法如下所述：

- **分析使用 oligo(dT) 作为引物合成的 cDNA 的长度**。在第一链 cDNA 的预实验中合成放射性标记的 cDNA。从哺乳动物 mRNA 合成的 cDNA 在电泳后应该呈现从约 600 bp 到大于 5 kb 的连续的扩散条带。放射性主要位于 1.5～2 kb，而且应该没有可见的 cDNA 条带，除非 mRNA 是从高度分化的、大量表达一组特定蛋白质的细胞制备而来的（如网状细胞和 B-淋巴细胞）。
- **在预实验的 northern 杂交中使用放射性标记的 poly（dT）作为探针**（Fornace and Mitchell 1986；Hollander and Fornace 1990）。Poly(A)$^+$ RNA 将会产生一个从 600 bp 到大于 5 kb 连续扩散的放射自显影图像。主要的放射活性位于 1.5～2 kb，并且仍然没有可见的特定 cDNA 条带，除非 mRNA 从表达大量特定的 mRNA 的细胞中制备。
- **使用 Northern 杂交检测从持家基因表达的已知长度的 mRNA**。例如，葡萄糖-6-磷酸脱氢酶（GAPDH）mRNA，在大多数哺乳动物细胞中期长度是 1.3 kb。条带显著地扩散为低分子质量的区域通常表明 mRNA 分子发生了明显的降解。

令人困惑的是，另一方面，这却是琼脂糖凝胶系统出现问题的一个标志（如在乙二醛中出现离子）。

272

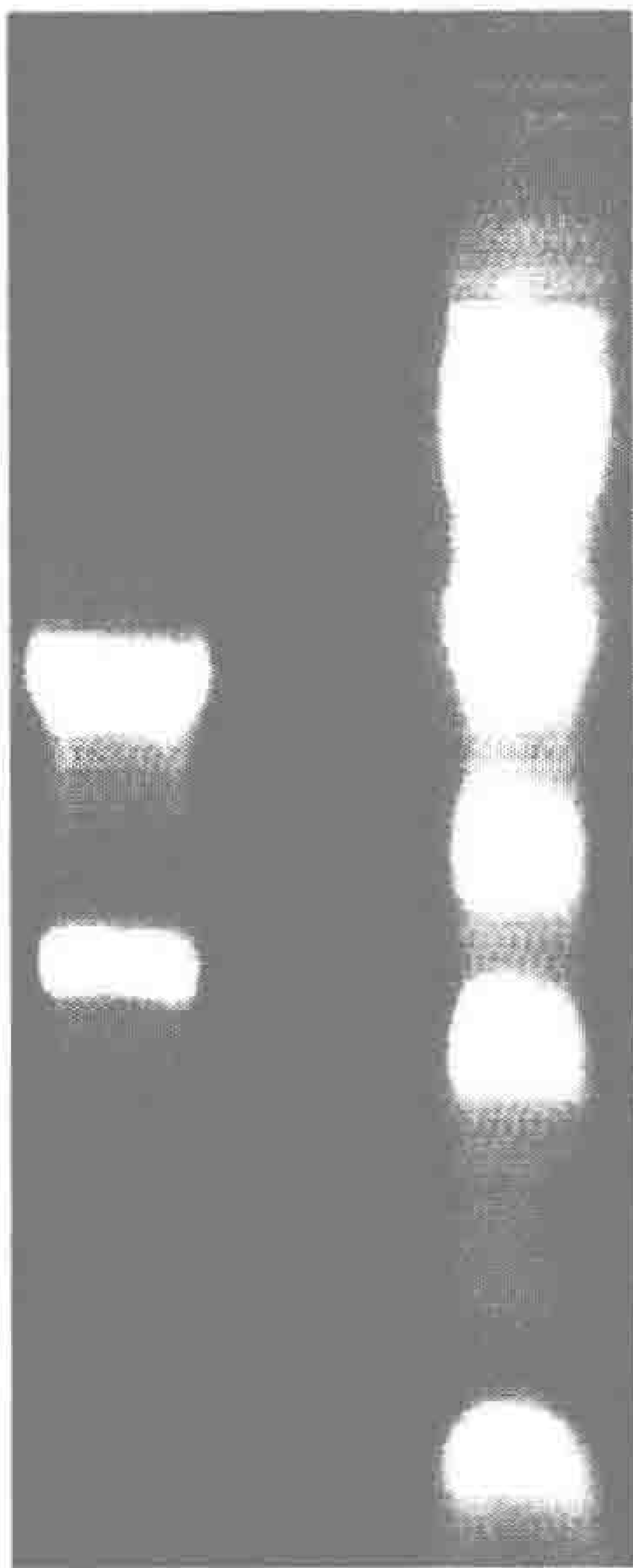


图 3-27 哺乳动物的 mRNA 在电泳之后，在溴化乙啶存在的情况下，rRNA 中的 28S 和 18S 种类在紫外光照射下应该清晰可见，与此同样的还应该有一条更加弥散的、快速迁移的条带，由 tRNA 和 5S rRNA 组成。如果 RNA 制备物未发生降解，28S rRNA 条带的染色强度应该是 18S 条带的大约两倍，并且这两个条带都不应该有可见的扩散。在加样孔附近有染色表明了制备物中仍然存在 DNA。除非凝胶过载，否则应该看不见mRNA（它扩散在凝胶中）

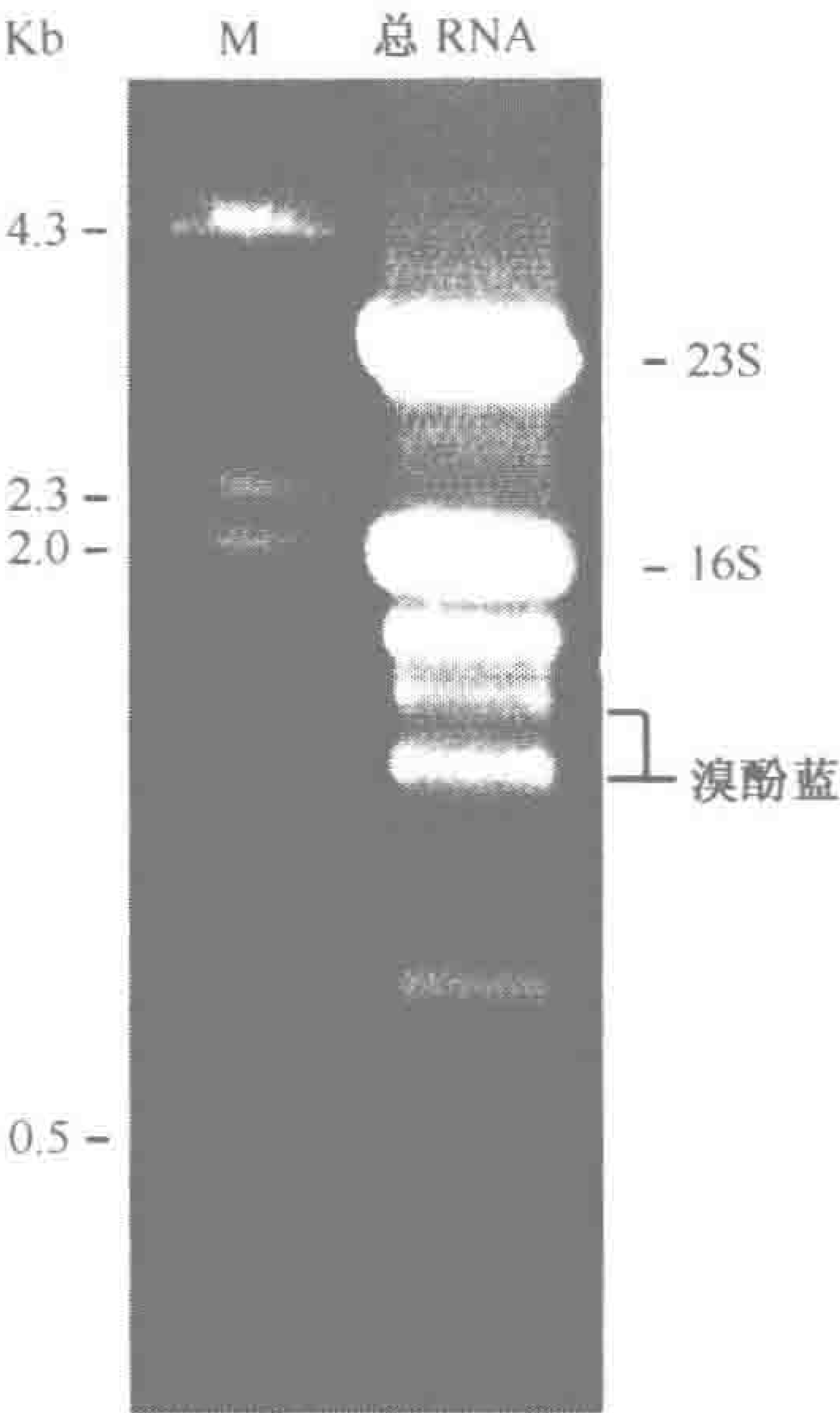


图 3-28 来自叶组织的总 RNA，其中富含叶绿体 RNA，通过 1.2%的琼脂糖电泳分离。高等植物的叶绿体包含有 70 S 核糖体，由一个大亚基和一个小亚基组成，分别是 50 S 和 30 S。大亚基含有相对分子质量为 1.1×10^6 的 23S rRNA，而小亚基含有相对分子质量为 0.56×10^6 的 16S rRNA(Kissil and Buetow 1982；Leaver 1982)。根据它们在核糖体中的化学定量关系，可以预计 23S RNA 要比 16 S 在重量上超出几乎两倍。然而，大多数 RNA 制备物中的 23S rRNA 的含量要比预期的少，最可能的原因是由于降解。因此，分离自植物的 rRNA 总是包含了降解产物，如图中所示。还要注意植物的核糖体含有一个短的 5S rRNA 分子，它在动物的线粒体中是没有的（Spremuli 2000）（照片由Bi-Huei Hou 提供，Carnegie Institute of Washington，Stanford University）

273

信息栏：测量 DNA 染料标记的效率

评价标记靶分子质量的最佳方法是琼脂糖凝胶电泳，它提供了关于靶分子的长度和它们的标记强度的信息（见方案 13）。然而，标记效率也可以通过标准的光谱方法测量，如这里所述。这种技术在估计应用于杂交反应的 Cy3-和 Cy5-标记的核苷酸数量时是非常有用的。

染料的标记效率，表示为碱基对染料的比率，可以这样进行估计：①测量核酸在 260 nm 的吸收值和染料在其最大吸收波长（λ）处的吸收值；②应用 Beer-Lambert 定律：

$$吸收值=消光系数\times光程长度\times浓度$$

常用于标记核酸的荧光染料的消光系数（ε）和最大吸收峰（λ_{max}）见表 3-11。

表 3-11　荧光染料的吸收特性

染　料	λ_{max}/nm	ε($cm^{-1}\cdot M^{-1}$)	修正因子(CF_{260})
Alexfluor 488	492(绿)	62 000	0.30
Alexfluor 532	525(黄)	82 300	0.24
Alexfluor 546	555(橙)	104 000	0.21
Alexfluor 568	576(红)	93 000	0.45
Alexfluor 594	588(红)	80 400	0.45
Cy3	550(橙)	150 000	0.08
Cy5	649(远红光)	250 000	0.05

1. 在 260 nm 和染料的 λ_{max}处测量标记 DNA 样品的吸收值（A）。要进行这些测量，溶液中应该含有大于 5 μ g/ml 的 DNA-染料结合物。
2. 修正染料对 A_{260}读数的贡献，使用修正因子 $A_{碱基}=A_{260}-(A_{染料}\times CF_{260})$。
3. 使用下面的公式，计算碱基与染料分子的比值。

碱基：染料＝$(A_{碱基}\times\varepsilon_{染料})/(A_{染料}\times\varepsilon_{碱基})$，式中，$\varepsilon_{染料}$是荧光染料的消光系数；$\varepsilon_{碱基}$是双链 DNA、单链 DNA 或者寡核苷酸中碱基的消光系数（见表 3-12）。

表 3-12　核酸的消光系数

274

核　酸	ε($cm^{-1}\cdot M^{-1}$)	MW_{base}
双链 DNA	6600	330
单链 DNA	8920	330
寡核苷酸	10 000	330

注：本表中的信息由 Amersham Biosciences，Molecular Probes，Inc.，以及 Ed Tom Lab 提供的文献汇总而来。

当使用 Cy3 或 Cy5 进行标记的时候，碱基：染料比值为 8～65，取决于模板 RNA 的特性（如 G+C）、含量和反转录酶的性质。由于 SUPERSCRIPT Ⅱ催化 Cy3 掺入的效率要比催化 Cy5 掺入更高，因此，Cy3 的标记密度通常要比 Cy5 标记高出 4/3 倍。

样品中的核酸浓度（C）可以用下面的公式计算：

$$C(mg/ml)=(A_{碱基}\times 330)/(\varepsilon_{碱基}\times 光程长度)$$

因为核酸的数量很少，通常使用小的比色池（容量为 100～200 μl）进行光谱测量。这些比色池的光程长度通常比常用的 1 cm 短。如果有可以测量光吸收的读板器，对紫外透明的微量滴定板也能使用。这些读板器的光程长度随型号和制造商的不同而变化，使用说明书中应该包含必需的数据。

信息栏：鼠白血病病毒反转录酶

Mo-MLV 反转录酶包含两个功能独立的、非重叠的区域（Tanese et al. 1985；Tanese and Goff 1988；Prasad 1993）。大的氨基末端区域（约 450 个残基）具有 DNA

聚合酶活性，而羧基末端区域（约 220 个残基）包含了 RNase H 活性。DNA 聚合酶非常不活跃并倾向于在 RNA 模板上富含二级结构的区域处停止（Matson et al. 1980；Gerard et al. 1989）。除此之外，RNase H 活性具有在链的起始延伸位置的附近水解 RNA 模板的能力，这将导致 DNA-RNA 杂交体的解离和第一链 cDNA 的终止。多达 50%的由 Mo-MLV 反转录酶开始合成的第一链 cDNA 分子可能会因此而过早地中止（Berger et al. 1983）。

275 在 20 世纪 90 年代早期，Mo-MLV 反转录酶三维空间结构的解决（Arnold et al. 1992；Kohlstaedt et al. 1992；Georgiadis et al. 1995）是一个重要的进步，因为它打开了对酶进行合理的遗传工程之门。然而，关键问题之一当时已经通过产生一个保留有 DNA 聚合酶活性而没有 RNase H 活性的 Mo-MLV 的突变体而得到解决（Kotewicz et al. 1988）。令人惊讶的是，通过缺失或者点突变使 RNase H 区域失活导致了酶在升高的温度下催化 cDNA 合成的能力极大提升（Gerard et al. 1989，1992）。在此之前，Mo-MLV一直在不超过 42℃的温度下使用。然而，缺乏 RNase H 活性的特定 Mo-MLV 突变体可以在高达 50℃的温度下有效地催化 cDNA 的合成（Gerard et al. 1989，1992），而 RNase H^- 形式的鸟类的酶能够在 60℃安全地使用（Gerard 1998）。（译者注：原文误印为 998）

从反转录酶消除 RNase H 活性增加了酶在标准温度下产生全长 cDNA 分子的能力。这些变异体的各种不同商品化形式包括一个缺乏 RNase H 活性并在羧基末端带有五个附加残基的 Mo-MLV 反转录酶突变体。其他缺乏 RNase H 活性的 Mo-MLV 反转录酶的商品化突变体形式有来自 Stratagene 的 StrataScripipt，来自 Invitrogen/GIBCO 的 SUPFRSCRIPT 和 SUPFRSCRIPT Ⅱ。SUPFRSCRIPT 带有一个消除 RNase H 活性的缺失，而 SUPFRSCRIPT 和 SUPFRSCRIPT Ⅱ包含了点突变。

信息栏：oligo-(dT) 纤维素

oligo-(dT) 仍旧被广泛用作分离和纯化 poly(A)$^+$ mRNA 的亲和配体，基本上与 Aviv 和 Leder(1972) 以及 Nakazato 和 Edmonds(1972) 所介绍的相同。为了这个目的，长度为 12～18 个残基的胸腺嘧啶寡聚体共价结合到固相的基质（通常是纤维素）上。寡聚体合成的标准方法由 Giham(1964，1971) 开创，包括了在碳二亚胺存在下胸腺嘧啶单磷酸盐的聚合。然后加入纤维素，寡聚胸腺嘧啶核苷酸分子通过它们的 5′-磷酰基与纤维素的羟基反应而结合上去。Giham(1964) 使用 oligo-(dT) 纤维素柱的目的主要是用于研究多聚核苷酸的亲和色谱，结果显示腺苷酸的各种寡聚体能够按照温度依赖性的方式从 oligo-(dT) 纤维素柱上洗脱下来。然而，对于 mRNA 的分离，更方便的是研究杂交反应对盐的依赖性，如方案 7 中介绍的。关于进一步的信息，见 Sambrook 和 Russell(2001)。

276

信息栏：RNA 的定量

由于吸收光谱法具有迅速、简单和非破坏性的特点，是长久以来用于在浓缩的纯溶

液中测量 DNA 和 RNA 含量的首选方法。表 3-13 显示了在 260 nm 处的光吸收与溶液中核酸的浓度之间的关系。

表 3-13　核酸在 260 nm 处的光吸收值

260 nm 处的光吸收值	消光系数	每单位 OD_{260} 核酸的量
RNA 和单链 DNA	7.4	40 μg/ml
双链 DNA	6.6	50 μg/ml
寡聚核苷酸	8.6	33 μg/ml

注：表中显示的数据适用于光程长度为 1 cm 的比色池；数据应该视作近似值，因为它们受到碱基的组成和单链核酸中二级结构水平的影响。

在单一波长（260 nm）测量其光吸收来估计核酸和寡核苷酸的浓度不是一种好方法。应该在多个波长处测量样品的光吸收，因为 260 nm 处的光吸收与其他波长处光吸收的比值能很好地反映制备物的纯度。在 230 nm 处有显著的吸收反映了受到酚盐离子、硫氰酸盐和其他有机化合物的污染（Stulnig and Amberger 1994），而在更高波长处的吸收（330 nm 或者更高）通常由光散射造成，并反映了存在着颗粒状物质。在 280 nm 处的吸收反映了蛋白质的存在。

多年以来，260 nm 和 280 nm 处光吸收的比值（OD_{260}/OD_{280}）被用来对分离核酸的纯度进行测量。该方法可以追溯到 1942 年，当时 Warburg 和 Christian 指出该比值是反映蛋白质制备物受到核酸污染的一个好的指标。倒过来使用就不正确了！因为核酸在 260 nm 和 280 nm 处的消光系数远远大于蛋白质的，蛋白质的严重污染不会使得核酸溶液的 OD_{260}/OD_{280} 比值发生很大的改变（Warburg and Christian 1942；Glasel 1995；Manchester 1995，1996；Wilfinger et al. 1997）。在理论上，RNA 的 A_{260}/A_{280} 比值应该为 1.9～2.0。然而，核酸的 A_{260}/A_{280} 比值受到溶液的 pH 和离子强度的显著影响（Beavan et al. 1955；Wilfinger et al. 1997）。在水中，当 pH 从 5.2 上升到 6.1 时，A_{260}/A_{280} 比值从 1.55 增高到 1.97。在 100 mmol/L NaCl 中，该比值为 1.5，而在 1 mol/L NaCl中，该比值为 2.1。受到大量酚污染的核酸制备物在 230 nm 处有强的吸收。

吸收光谱法很简单，但灵敏度相对较低。对于大多数实验室中使用的分光光度计，为了得到对 A_{260} 可靠的估计，核酸浓度至少要为 1 μg/ml。此外，吸收光谱法不能区分 DNA 和 RNA，并且不适用于核酸的粗制品。由于有这些限制，发明了许多其他方法用于测量 RNA 的浓度。

RiboGreen

277

RiboGreen 是由 Molecular Probes Inc. 销售的一种非对称的青色素染料。游离的染料基本上是没有荧光的，但是与 RNA 结合以后在 520 nm 处显示了 1000 倍的荧光增强。可以在分光光度计、读板器或者配备了标准荧光滤色镜而基于滤色镜的荧光光度计（如 Turner Designs TD-700）中测量低至 10 ng/ml 的 RNA。RiboGreen 对 RNA 的灵敏度比溴化乙啶高 200 倍，并且比 260 nm 处的光吸收高 100 倍。分析时需要两种浓度

的染料以覆盖它的全部动态范围（10 ng/ml～1 mg/ml）。主要的缺点是：①费用高；②染料不能区分 RNA 和 DNA。可以在分析前用 DNase 处理样品来解决这个问题。测试样品中 RNA 的浓度，通过与使用已知浓度的 RNA 样品建立的标准曲线进行对比来确定。关于分析的详细情况可以从 Molecular Probes 网站（www. molecularprobes. com）得到，同时在 RiboGreen 试剂盒提供的文献内也可以找到。

信息栏：RNA 操作的特别注意事项

在实验中很容易会对 RNase 的污染放不下心，这是不必要的。与内源的 RNase 有关的问题通常可以通过对组织的正确处理而消除。理想情况下，组织在液氮中冷冻、保存在 RNA*later* 中或者迅速在 TRIzol 中匀浆。关于更详细的情况，见表 3-1。

对 RNA 来说，最大的风险是组织分离的时间和在变性溶液中冷冻或者裂解的时间。为了使内源的 RNase 不至于破坏 RNA，缩短这个时间和避免损伤组织是非常关键的。如果不能够立即抽提 RNA，将组织保存在 RNA*later*(Ambion) 中（表 3-1）。

在纯化过程接近完成的时候也需要特别注意，这时 RNA 转移到水溶液中，失去了酚和胍离子等试剂提供的保护。在此阶段，如果 Rnase 没有被除去或者失活，或者引入了 RNase，即使是很微量的，也会降解 RNA。操作 RNA 时动作要保持迅速，缩短在大于 0℃条件下花费的时间（见关于 **RNA 的保存和恢复**信息栏）。

在我们的实验中，外源 RNase 的污染常常是由于无菌技术的失败，溶液遭到细菌或其他微生物的污染，或者来自以前用于分配 RNase 溶液的自动移液装置，如在质粒的制备过程中。下面列举的是避免 RNase 污染的有效方法。

- 保留一套专用的自动移液器用于 RNA 的操作。
- 将专门用在和 RNA 相关实验的玻璃器皿、成批的塑料制品和缓冲液单独存放。作为替代，也可以使用一次性的塑料制品，这些在使用前通常不需要清洁。可以得到保证没有 RNase 的一次性的移液器吸头和微量离心管。为了减少污染的机会，将这些小的物品从它们原始的包装转移到实验室的架子上时，最好使用灭菌的镊子。

- 将溶液/缓冲液分装成小份保存，并在使用后将每份剩余的丢弃。避免实验材料或者储存液用于实验室中其他的用途。
- 永远不要假定实验室中的任何物品是不含 RNase 的，除非亲自进行了下述除去或者灭活酶的操作步骤。
- 准备专用的电泳槽用于 RNA 的凝胶电泳。用去污剂清洗这些电泳槽，在水中清洗，用乙醇干燥，然后向其中注满 3%的 H_2O_2溶液。室温下放置 10 min 后，用经过 DEPC 处理的 H_2O 彻底清洗电泳槽。含有伯胺的溶液（如 Tris）不能用 DEPC 处理（见关于**抑制 RNase** 的信息栏）。
- 所有溶液用不含 RNase 的玻璃器皿、H_2O 和留作 RNA 实验专用并经过烘焙（300℃处理 4 h）的刮勺取用的化学试剂配制。如果条件允许，在 37℃下用 0.1%的 DEPC 处理溶液至少 12 h，然后再在 100℃下加热 15 min，或者在 15 psi(1.05 kg/cm²）下使用液体灭菌程序高压灭菌 15 min（关于进一步的信

息，见关于**抑制 RNase** 的信息栏)。热不稳定的溶液用 0.2 μm 的过滤器过滤，并保存在不含 RNase 的玻璃或者塑料器皿中。

- 操作称量皿或者处理 RNA 的时候要戴上手套。要经常更换手套。
- 经过高压灭菌的玻璃器皿可能不足以使 RNase 失活。将全部玻璃器皿在 300℃下烘焙 4 h，或者注满 0.5 mol/L NaOH 并放置 30 min，然后用灭菌的、不含 RNase 的 H_2O 清洗数次。
- 将微量离心管在 0.1%的 DEPC 中浸泡过夜，然后用灭菌的，不含 RNase 的 H_2O 清洗并高压灭菌，使它们不带有 RNase。
- 使用前用 RNase-Off(CPG Inc.) 或者其他去除 RNase 污染的试剂擦洗来清除公用仪器设备，如旋转蒸发器的污染。

信息栏：RNA 的保存和回收

RNA 的保存

在乙醇沉淀后，按照下述的方法保存 RNA：

- 将沉淀物溶解于去离子的甲酰胺并保存在－20℃(Chomczynski 1992)。这提供了一个稳定的化学环境，同时也防止了 RNA 被 RNase 降解。经过纯化的，不含盐的 RNA 在甲酰胺中迅速溶解，最高浓度可达 4 mg/ml。在这样的浓度下，RNA 样品可以直接使用凝胶电泳和 RT-PCR 分析，或者用作由反转录酶催化的 cDNA 合成的模板，因此节约了时间并防止了可能的降解。如果需要的话，按照 Chomczynski(1992) 介绍的方法用 4 倍体积的乙醇进行沉淀从甲酰胺中回收 RNA，或者用 0.2 mol/L 的 NaCl 将甲酰胺稀释 4 倍，然后加入通常的 2 倍体积的乙醇 (Nadin-Davis and Mezl 1982)。
- 将沉淀物溶解于缓冲的水溶液中并保存在－80℃。这里常用的缓冲液有 TE (pH 7.4) 或者不含 RNase 的 H_2O，其中加入 0.1 mol/L 的含有 SDS(0.1%～0.5%) 的 EDTA(pH 7.5)。该方法的一个缺点是在 RNA 用作反转录酶的模板之前，SDS 必须通过氯仿抽提和乙醇沉淀除去。而且，在使用分装的小份样品之前，必须进行加温，这会使 RNA 遭到残留的 RNase 的消化。
- 将 RNA 沉淀物以乙醇悬液的形式保存在－20℃。RNA 样品可以根据需要用自动移液装置移取。将 RNA 储存在乙醇中的优点有：①具有低的冰点，可使小份 RNA 的回收能够在样品加温到大于 0℃的温度下进行；②RNase 在乙醇中是没有活性的。然而，由于 RNA 沉淀是块状的，在取样前要立即将悬液涡旋振荡处理。

从保存物中回收 RNA

- RNA 通常保存在甲酰胺、乙醇或者含有 SDS 的 Tris/EDTA 溶液中。
- 按照 Chomczynski(1992) 介绍的方法，用 4 倍体积的乙醇沉淀保存在甲酰胺中的 RNA，然后进行步骤 1。
- 加入 0.1 倍体积的 3 mol/L 醋酸钠 (pH 5.2) 和 3 倍体积的无水乙醇以沉淀保

存在 TE/SDS 中的 RNA，继续步骤 1。

- 按照步骤 1 中的介绍，通过离心回收保存在乙醇中的 RNA 沉淀。

1. 4℃下，以 12 000 *g* 离心 10 min，回收 RNA。

   如果大量的 RNA(>150 μg) 在离心管中长时间离心，会形成一个致密的、坚硬的团块，而不容易溶解。在处理大量的 RNA 时，将样品离心 30 s 以回收大部分的样品。将大部分的上清液转移到一支新的离心管中，然后再在 12 000 *g* 下继续离心 10 min 回收残留的 RNA。在下面的三步操作中，将这两批的 RNA 分开处理，然后在步骤 5 将它们合并。

2. 轻轻地吸出上清液，然后加入 1 ml 75％的乙醇。倒转密封的离心管数次，然后在 4℃下，以 12 000 *g* 再次离心 5 min。

280

3. 小心倾去上清液，将离心管再次短暂离心并用自动移液器除去残留的乙醇。
4. 将开口的试管竖立在通风橱中直到 RNA 刚好变干，时间不要过长。
5. 将 RNA 重悬于 50 μl 10 mmol/L Tris-Cl(pH 7.0)、0.1 mmol/L EDTA。

   用于微阵列实验，RNA 要浓缩到大于 7 mg/ml。

6. 将密封的试管在 65℃下温育 5～10 min。将试管短暂涡旋使液体集中在管底。

   在缓冲的水溶液中，RNA 团块迅速变得透明，因此，摇晃溶液检查 RNA 是否重新溶解的时候，必须将样品拿到光下。

7. 测量在 260 nm 和 280 nm 的光吸收对 RNA 定量，或者用 RiboGreen 染色后通过荧光定量（见关于 **RNA 的定量**信息栏）。

   将 RNA 用于微阵列实验之前，通过甲醛-1％琼脂糖凝胶电泳检查其完整性（见关于**哺乳动物、植物和细菌 RNA** 的信息栏）。

参考文献

281

Aharoni A., Keizer L.C.P., Bouwmeester H.J., Sun Z., Alvarez-Huerta M., Verhoeven H.A., Blaas J., van Houwelingen A.M.M.L., De Vos R.C.H., van der Voet H., Jansen R.C., Guis M., Mol J., Davis R.W., Schena M., van Tunen A.J., and O'Connell A.P. 2000. Identification of the SAAT gene involved in strawberry flavor biogenesis by use of DNA microarrays. *Plant Cell* **12:** 647–661.

Arnold E., Jacobo-Molina A., Nanni R.G., Williams R.L., Lu X., Ding J., Clark Jr., A.D., Zhang A., Ferris A.L., Clark P., et al. 1992. Structure of HIV-1 reverse transcriptase/DNA complex at 7 Å resolution showing active site locations. *Nature* **357:** 85–89.

Aviv H. and Leder H. 1972. Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid cellulose. *Proc. Natl. Acad. Sci.* **69:** 1408–1412.

Barlow J.J., Mathias A.P., and Williamson R. 1963. A simple method for the quantitative isolation of undegraded high molecular weight ribonucleic acid. *Biochem. Biophys. Res. Commun.* **13:** 61–66.

Barnes W.M. 1994. PCR amplification of up to 35-kb DNA with high fidelity and high yield from lambda bacteriophage templates. *Proc. Natl. Acad. Sci.* **91:** 2216–2220.

Baugh L.R., Hill A.A., Brown E.L., and Hunter C.P. 2001. Quantitative analysis of mRNA amplification by in vitro transcription. *Nucleic Acids Res.* **29:** E29.

Beavan G.H., Holiday E.R., and Johnson E.A. 1955. Optical properties of nucleic acids and their components. In *The nucleic acids: Chemistry and biology* (ed. E. Chargaff and J.N. Davidson), vol. 1, pp. 493–553. Academic Press, New York.

Berger S.L. and Birkenmeier C.S. 1979. Inhibition of intractable nucleases with ribonucleotide-vanadyl complexes: Isolation of messenger ribonucleic acid from resting lymphocytes. *Biochemistry* **18:** 5143–5149.

Berger S.L., Wallace D.M., Puskas R.S., and Eschenfeldt W.H. 1983. Reverse transcriptase and its associated ribonuclease H: Interplay of two enzyme activities controls the yield of single-stranded complementary deoxyribonucleic acid. *Biochemistry* **22:** 2365–2372.

Berk A.J. and Sharp P.A. 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. *Cell* **12:** 721–732.

Bertucci F., Bernard K., Loriod B., Chang Y.C., Granjeaud S., Birnbaum D., Nguyen C., Peck K., and Jordan B.R. 1999. Sensitivity issues in DNA array-based expression measurements and performance of nylon microarrays for small samples. *Hum. Mol. Genet.* **8:** 1715–1722.

Bishop J.O., Morton J.G., Rosbash M., and Richardson M. 1974. Three abundance classes in HeLa cell messenger RNA. *Nature* **250:** 199–204.

Blackburn P., Wilson G., and Moore S. 1977. Ribonuclease inhibitor from human placenta. Purification and properties. *J. Biol.Chem.* **252:** 5904–5910.

Blattner F.R., Plunkett III G., Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., and Shao Y. 1997. The complete genome sequence of *Escherichia coli* K-12. *Science* **277:** 1453–1474.

Bodkin D.K. and Knudson D.L. 1985. Assessment of sequence relatedness of double-stranded RNA genes by RNA-RNA blot hybridization. *J. Virol. Methods* **10:** 45–52.

Bolton E.T. and McCarthy B.J. 1962. A general method for isolation of RNA complementary to DNA. *Proc. Natl. Acad. Sci.* **48:** 1390–1396.

Bonner J., Kung G., and Bekhor I. 1967. A method for the hybridization of nucleic acid molecules at low temperature. *Biochemistry* **6:** 3650–3653.

Brawerman G., Mendecki J., and Lee S.Y. 1972. A procedure for the isolation of mammalian messenger ribonucleic acid. *Biochemistry* **11:** 637–641.

Brown J.W.S. and Shaw P.J. 1998. Small nucleolar RNAs and pre-rRNA processing in plants. *Plant Cell* **10:** 649–657.

Brown N.L. 1984. DNA sequencing. *Methods Microbiol.* **17:** 259–313.

Brumbaugh J.A., Middendorf L.R., Grone D.L., and Ruth J.L. 1988. Continuous, on-line DNA sequencing using oligodeoxynucleotide primers with multiple fluorophores. *Proc. Natl. Acad. Sci.* **85:** 5610–5614.

Casey J. and Davidson N. 1977. Rates of formation and thermal stabilities of RNA:DNA and DNA:DNA duplexes at high concentrations of formamide. *Nucleic Acids Res.* **4:** 1539–1552.

Celano P., Vertino P.M., and Casero Jr. R.A. 1993. Isolation of polyadenylated RNA from cultured cells and intact tissues. *BioTechniques* **15:** 26–28.

Chang S., Puryear Y., and Cairney J. 1993. A simple and efficient method for isolating RNA from pine trees. *Plant Mol. Biol. Rep.* **11:** 113–116.

Chen J.J., Wu R., Yang P.C., Huang J.Y., Sher Y.P., Han M.H., Kao W.C., Lee P.J., Chiu T.F., Chang F., Chu Y.W., Wu C.W., and Peck K. 1998. Profiling expression patterns and isolating differentially expressed genes by cDNA microarray system with colorimetry detection. *Genomics* **51:** 313–324.

Chenchik A., Zhu Y.Y., Diatchenko L., Li R., Hill J., and Siebert P.D. 1998. Generation and use of high-quality cDNA from small amounts of total RNA by SMART™ PCR. In *Gene cloning and analysis by RT-PCR* (ed. P.D. Siebert and J.W. Larrick), pp. 305–319. Eaton Publishing/BioTechniques Book Division, Natick, Massachusetts.

282

Chirgwin J.M., Przybyla A.E., MacDonald R.J., and Rutter W.J. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* **18:** 5294–5299.

Chomczynski P. 1989. Product and process for isolating RNA. U.S. Patent 4,843,155.

———. 1992. One-hour downward alkaline capillary transfer for blotting of DNA and RNA. *Anal. Biochem.* **201:** 134–139.

———. 1993. A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. *BioTechniques* **15:** 532–534.

———. 1994. Shelf-stable product and process for isolating RNA, DNA and proteins. U.S. Patent 5,346,994.

———. 1999. Product and process for isolating DNA, RNA and proteins. U.S. Patent 5,945,515.

Chomczynski P. and Mackey K. 1995. Modification of the TRI Reagent™ procedure for isolation of RNA from polysaccharide- and proteoglycan-rich sources. *BioTechniques* **19:** 942–945.

Chomczynski P. and Sacchi N. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **221:** 303–305.

Chuang S.E., Daniels D.L., and Blattner F.R. 1993. Global regulation of gene expression in *Escherichia coli. J. Bacteriol.* **175:** 2026–2036.

Cox R.A. 1968. The use of guanidinium chloride in the isolation of nucleic acids. *Methods Enzymol.* **12B:** 120–129.

Crino P.B. and Eberwine J. 1997. Cellular and molecular basis of cerebral dysgenesis. *J. Neurosci. Res.* **50:** 907–916.

Defrancesco L. 1998. Need for speed II. *The Scientist* **12:** 20–26.

Deyholos M.K. and Galbraith D.W. 2001. High-density microarrays for gene expression analysis. *Cytometry* **43:** 229–238.

DeRisi J.L., Iyer V.R., and Brown P.O. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278:** 680–686.

DeRisi J., Penland L., Brown P.O., Bittner M.L., Meltzer P.S., Ray M., Chen Y., Su Y.A., and Trent J.M. 1996. Use of a cDNA microarray to analyze gene expression patterns in human cancer. *Nat. Genet.* **14:** 457–460.

Diehn M., Eisen M.B., Botstein D., and Brown P.O. 2000. Large-scale identification of secreted and membrane-associated gene products using DNA microarrays. *Nat. Genet.* **25:** 58–62.

Duggan D.J., Bittner M., Chen Y., Meltzer P. and Trent P.M. 1999. Expression profiling using cDNA microarrays. *Nat. Genet.* (suppl.) **21:** 10–14.

Dunn J.J. and Studier F.W. 1983. Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J. Mol. Biol.* **166:** 477–535.

Eberwine J., Yeh H., Miyashiro K., Cao Y., Nair S., Finnell R., Zettel M., and Coleman P. 1992. Analysis of gene expression in single live neurons. *Proc. Natl. Acad. Sci.* **89:** 3010–3014.

Edmonds M., Vaughn Jr. M.H. and Nakazato H. 1971. Polyadenylic acid sequences in the heterogeneous nuclear RNA and rapidly-labeled polyribosomal RNA of HeLa cells: Possible evidence for a precursor relationship. *Proc. Natl. Acad. Sci.* **68:**1336–1340.

Ehrenberg L., Fedorcsak I., and Solymosy F. 1976. Diethyl pyrocarbonate in nucleic acid research. *Prog. Nucleic Acid Res. Mol. Biol.* **16:** 189–192.

Eisen M.B. and Brown P.O. 1999. DNA arrays for analysis of gene expression. *Methods Enzymol.* **303:** 179–205.

Florell S.R., Coffin C.M., Holden J.A., Zimmermann J.W., Gerwels J.W., Summers B.K., Jones D.A., and Leachman S.A. 2001. Preservation of RNA for functional genomic studies: A multidisciplinary tumor bank protocol. *Mod. Pathol.* **14:** 116–128.

Fornace Jr. A.J. and Mitchell J.B. 1986. Induction of B2 RNA polymerase III transcription by heat shock enrichment for heat shock induced sequences in rodent cells by hybridization subtraction. *Nucleic Acids Res.* **14:** 5793–5811.

Freimert C., Erfle V., and Strauss G. 1989. Preparation of radiolabeled cDNA probes with high specific activity for rapid screening of gene expression. *Methods Mol. Cell. Biol.* **1:** 143–153.

Gaillard C. and Strauss F. 1990. Ethanol precipitation of DNA with linear polyacrylamide as carrier. *Nucleic Acids Res.* **18:** 378.

Gasch A.P., Spellman P.T., Kao C.M., Carmel-Harel O., Eisen M.B., Storz G., Botstein D., and Brown P.O. 2000. Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell* **11:** 4241–4257.

Georgiadis M.M., Jessen S.M., Ogata C.M., Telesnitsky A., Goff S.P., and Hendrickson W.A. 1995. Mechanistic implications from the structure of a catalytic fragment of Moloney murine leukemia virus reverse transcriptase. *Structure* **3:** 879–892.

Gerard G.F. 1998. Reverse transcriptase: A historical perspective. *Focus* (Life Technologies) **20:** 65–67.

283

Gerard G.F., D'Alessio J.M., and Kotewicz M.L. 1989. cDNA synthesis by cloned Moloney murine leukemia virus reverse transcriptase lacking RNase H activity. *Focus* (Life Technologies) **11:** 66–69.

Gerard G.F., Schmidt B.J., Kotewicz M.L., and Campbell J.H. 1992. cDNA synthesis by Moloney murine leukemia virus RNase H-minus reverse transcriptase possessing full DNA polymerase activity. *Focus* (Life Technologies) **14:** 91–93.

Getz M.J., Birnie G.D., Getz M.J., Birnie G.D., Young B.D., MacPhail E., and Paul J. 1975. A kinetic estimation of base sequence complexity of nuclear poly(A)-containing RNA in mouse Friend cells. *Cell* **4:** 121–129.

Getz M.J., Reiman Jr. H.M., Siegal G.P., Quinlan T.J., Proper J., Elder P.K., and Moses H.L. 1977. Gene expression in chemically transformed mouse embryo cells: Selective enhancement of the expression of C type RNA tumor virus genes. *Cell* **11:** 909–921.

Gilham P.T. 1964. Synthesis of polynucleotide celluloses and their use in fractionation of oligonucleotides. *J. Am. Chem. Soc.* **86:** 4982–4989.

———. 1971. The covalent binding of nucleotides, polynucleotides and nucleic acids to cellulose. *Methods Enzymol.* **21:** 191–197.

Glasel J.A. 1995. Validity of nucleic acid purities monitored by 260nm/280nm absorbance ratios (comments). *BioTechniques* **18:** 62–63.

Glisin V., Crkvenjakov R., and Byus C. 1974. Ribonucleic acid isolated by cesium chloride centrifugation. *Biochemistry* **13:** 2633–2637.

Goldberg R.B., Hoschek G., Kamalay J.C., and Timberlake W.E. 1978. Sequence complexity of nuclear and polysomal RNA in leaves of the tobacco plant. *Cell* **14:** 123–131.

Gramza A.W., Lucas J.M., Mountain R.E., Schuller D.E., and Lang J.C. 1995. Efficient method for preparing normal and tumor tissue for RNA extraction. *BioTechniques* **18:** 228–231.

Grassman W. and Deffner G. 1953. Verteilungschromatographisches verhalten von proteinen un peptiden in phenolhaltigen lösungsmitteln. *Hoppe-Seyler's Z. Physiol. Chem.* **293:** 89–98.

Gress T.M., Hoheisel J.D., Lennon G.G., Zehetner G., and Lehrach H. 1992. Hybridization fingerprinting of high-density cDNA-library arrays with cDNA pools derived from whole tissues. *Mamm. Genome* **3:** 609–619.

Groppe J.C. and Morse D.E. 1993. Isolation of full-length RNA templates for reverse transcription from tissues rich in RNase and proteoglycans. *Anal. Biochem.* **210:** 337–343.

Guatelli J.C., Whitfield K.M., Kwoh D.Y., Barringer K.J., Richman D.D., and Gingeras T.R. 1990. Isothermal, in vitro amplification of nucleic acids by a multienzyme reaction modeled after retroviral replication. *Proc. Natl. Acad. Sci.* **87:** 7797.

Gubler U. and Hoffman B.J. 1983. A simple and very efficient method for generating cDNA libraries. *Gene* **25:** 263–269.

Hamaguchi K. and Geiduschek P. 1962. The effect of electrolytes on the stability of the deoxynucleate helix. *J. Am. Chem. Soc.* **84:** 1329–1338.

Harmer S.L., Hogenesch J.B., Straume M., Chang H.S., Han B., Zhu T., Wang X., Kreps J.A., and Kay S.A. 2000. Orchestrated transcription of key pathways in *Arabidopsis* by the circadian clock. *Science* **290:** 2110–2113.

Hearst J.E. and Vinograd J. 1961a. The net hydration of deoxyribonucleic acid. *Proc. Natl. Acad. Sci.* **47:** 825–830.

———. 1961b. The net hydration of T4 bacteriophage deoxyribonucleic acid and the effect of hydration on buoyant behavior in a density gradient at equilibrium in the ultracentrifuge. *Proc. Natl. Acad. Sci.* **47:** 1005–1014.

Heller R.A., Schena M., Chai A., Shalon D., Bedilion T., Gilmore J., Woolley D.E., and Davis R.W. 1997. Discovery and analysis of inflammatory disease-related genes using cDNA microarrays. *Proc. Natl. Acad. Sci.* **94:** 2150–2155.

Helmkamp G. and Ts'o P.O.P. 1961. The secondary structures of nucleic acids in organic solvents. *J. Am. Chem. Soc.* **83:** 138–142.

Hollander M.C. and Fornace Jr. A.J. 1990. Estimation of relative mRNA content by filter hybridization to a polythymidylate probe. *BioTechniques* **9:** 174–179.

Huang Y.H., Leblanc P., Apostolou V., Stewart B., and Moreland R.B. 1995. Comparison of Milli-Q® PF Plus water to DEPC-treated water in the preparation and analysis of RNA. *BioTechniques* **19:** 656–661.

Hughes T.R., Mao M., Jones A.R., Burchard J., Marton M.J., Shannon K.W., Lefkowitz S.M., Ziman M., Schelter J.M., Meyer M.R., Kobayashi S., Davis C., Dai H., He Y.D., Stephaniants S.B., Cavet G., Walker W.L., West A., Coffey E., Shoemaker D.D., Stoughton R., Blanchard A.P., Friend S.H., and Linsley P.S. 2001. Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. *Nat. Biotechnol.* **19:** 342–347.

Hutton J.R. 1977. Renaturation kinetics and thermal stability of DNA in aqueous solutions of formamide and urea. *Nucleic Acids Res.* **4:** 3537–3555.

284

Irving J.A., Cain G., Parr A., Howard M., Angus B., and Cattan A.R. 1996. OCT embedded sections of pathological specimens as a source of high quality RNA for reverse transcriptase/polymerase chain reaction. *J. Clin. Pathol.* **49:** 258–259.

Kacharmina J.E., Crino P.B., and Eberwine J. 1999. Preparation of cDNA from single cells and subcellular regions. *Methods Enzymol.* **303:** 3–18.

Karrer E.E., Lincoln J.W., Hogenhout S., Bennett A.B., Bostock R.M., Martineau B., Lucas W.J., Gilchrist D.G., and Alexander D. 1995. *In situ* isolation of mRNA from individual plant cells: Creation of cell-specific cDNA libraries. *Proc. Natl. Acad. Sci.* **92:** 3814–3818.

Katterman F.R.H. and Shattuck V.I. 1983. An effective method of DNA isolation from the mature leaves of *Gossipium* species that contain large amounts of phenolic terpenoids and tannins. *Prep. Biochem.* **13:** 347–359.

Kehoe D.M., Villand P., and Somerville S. 1999. DNA microarrays for studies of higher plants and other photosynthetic organisms. *Trends Plant Sci.* **4:** 38–41.

Khodursky A.B., Peter B.J., Cozzarelli N.R., Botstein D., Brown P.O., and Yanofsky C. 2000. DNA microarray analysis of gene expression in response to physiological and genetic changes that affect tryptophan metabolism in *Escherichia coli. Proc. Natl. Acad. Sci.* **97:** 12170–12175.

Kirby K.S. 1956. A new method for the isolation of ribonucleic acid from mammalian tissues. *Biochem. J.* **64:** 405–408.

———. 1957. A new method for the isolation of deoxyribonucleic acids; evidence on the nature of bonds between deoxyribonucleic acid and protein. *Biochem. J.* **66:** 495–504.

———. 1964. Isolation and fractionation of nucleic acids. *Prog. Nucleic Acid Res. Mol. Biol.* **3:** 1–31.

Kissil M.S. and Buetow D.E. 1982. Preparation of in vivo labelled, high specific activity [^{32}P]chloroplast rRNA from *Euglena gracilis*. In *Methods in chloroplast molecular biology* (ed. M. Edellman et al.), pp. 377–403. Elsevier Biomedical Press, New York.

Kohlstaedt L.A., Wang J., Friedman J.M., Rice P.A., and Steitz T.A. 1992. Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* **256:** 1783–1790.

Köhrer K. and Domdey H. 1991. Preparation of high molecular weight RNA. *Methods Enzymol.* **194:** 398–405.

Kondo T., Nakajima Y., and Kawakami M. 1979. Effect of salts and chromatin concentrations on the buoyant density of chromatin in metrizamide gradient. *Biochim Biophys Acta* **561:** 526–534.

Kotewicz M.L., Sampson C.M., D'Alessio J.M., and Gerard G.F. 1988. Isolation of cloned Moloney murine leukemia virus reverse transcriptase lacking ribonuclease H activity. *Nucleic Acids Res.* **16:** 265–277.

Lanotte M., Martin-Thouvenin V., Najman S., Balerini P., Valensi F., and Berger R. 1991. NB4, a maturation inducible cell line with t(15;17) marker isolated from a human acute promyelocytic leukemia (M3). *Blood* **77:** 1080–1086.

Leaver C.J. 1982. Isolation of intact 16S and 23S ribosomal RNAs from higher plants. In *Methods in chloroplast molecular biology* (ed. M. Edellman et al.), pp. 369–375. Elsevier Biomedical Press, New York.

Lee F.S. and Vallee B.L. 1993. Structure and action of mammalian ribonuclease (angiogenin) inhibitor. *Prog. Nucleic Acid Res. Mol. Biol.* **44:** 1–30.

Lee F.S., Shapiro R., and Vallee B.L. 1989. Tight binding inhibition of angiogenin and ribonuclease A by placental ribonuclease inhibitor. *Biochemistry* **28:** 225–230.

Lehrach H., Diamond D., Wozney J.M., and Boedtker H. 1977. RNA molecular weight determinations by gel electrophoresis under denaturing conditions: A critical reexamination. *Biochemistry* **16:** 4743–4751.

Lennon G.G. and Lehrach H. 1991. Hybridization analyses of arrayed cDNA libraries. *Trends Genet.* **7:** 314–317.

Lipshultz R.J., Fodor S.P.A., Gingeras T.R., and Lockhart D.J. 1999. High density synthetic oligonucleotides arrays. *Nat. Genet.* **21:** 20–24.

Litt M., Hauge X., and Sharma V. 1993. Shadow bands seen when typing polymorphic dinucleotide repeats: Some causes and cures. *BioTechniques* **15:** 280–284.

Lockhart D.J., Dong H., Byrne M.C., Follettie M.T., Gallo M.V., Chee M.S., Mittmann M., Wang C., Kobayashi M., Horton H., and Brown E.L. 1996. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Biotechnol.* **14:** 1675–1680.

Luo L., Salunga R.C., Guo H., Bittner A., Joy K.C., Galindo J.E., Xiao H., Rogers K.E., Wan J.S., Jackson M.R., and Erlander M.G. 1999. Gene expression profiles of laser-captured adjacent neuronal subtypes. *Nat. Med.* **5:** 117–122.

Maier E., Meier-Ewart S., Ahmadi A.R., Curtis J., and Lehrach H. 1994. Application of robotic technology to automated sequence fingerprint analysis by oligonucleotide hybridisation. *J. Biotechnol.* **35:** 191–203.

Maleck K., Levine A., Eulgen T., Morgan A., Schmid J., Lawton K.A., Dangl J.L., and Dietrich R.A. 2000. The transcriptome of *Arabidopsis* during systemic acquired resistance. *Nat. Genet.* **26:** 403–410.

285

Manchester K.L. 1995. Value of A_{260}/A_{280} ratios for measurement of purity of nucleic acids. *BioTechniques* **19:** 208–210.

———. 1996. Use of UV methods for measurement of protein and nucleic acid concentrations. *BioTechniques* **20:** 968–970.

Marmur J. and T'so P.O.P. 1961. Denaturation of deoxyribonucleic acid by formamide. *Biochim. Biophys. Acta* **51:** 32–36.

Martin R. 1987. Overcoming DNA sequencing artifacts: Stops and compressions. *Focus* (Life Technologies) **9:** 8–10.

Matson S.W., Fay P.J., and Bambara R.A. 1980. Mechanism of inhibition of the avian myeloblastosis virus deoxyribonucleic acid polymerase by adriamycin. *Biochemistry* **19:** 2089–2096.

Matz M., Shagin D., Bogdanova E., Britanova O., Lukyanov S., Diatchenko L., and Chenchik A. 1999. Amplification of cDNA ends based on template-switching effect and step-out PCR. *Nucleic Acids Res.* **27:** 1558–1560.

McConaughy B.L., Laird C.D., and McCarthy B.J. 1969. Nucleic acid reassociation in formamide. *Biochemistry* **8:** 3289–3295.

McGonigle B., Keeler S.J., Lau S.MC., Koeppe M.K., and O'Keefe D.P. 2000. A genomics approach to the comprehensive analysis of the glutathione *S*-transferase gene family in soybean and maize. *Plant Physiol.* **124:** 1105–1120.

Mechler B.M. 1987. Isolation of messenger RNA from membrane-bound polysomes. *Methods Enzymol.* **152:** 241–248.

Mujumdar R.B., Ernst L.A., Mujumdar S.R., Lewis C.J., and Waggoner A.S. 1993. Cyanine dye labeling reagents: Sulfoindocyanine succinimidyl esters. *Bioconjug. Chem.* **4:** 105–111.

Murphy N.R., Leinbach S.S., and Hellwig R.J. 1995. A potent, cost-effective RNAse inhibitor. *BioTechniques* **18:** 1068–1073.

Murray A.E., Lies D., Li G., Nealson K., Zhou J., and Tiedje J.M. 2001. DNA/DNA hybridization to microarrays reveals gene-specific differences between closely related microbial genomes. *Proc. Natl. Acad. Sci.* **98:** 9853–9858.

Nadin-Davis S. and Mezl V.A. 1982. Optimization of the ethanol precipitation of RNA from formamide containing solutions. *Prep. Biochem.***12:** 49–56.

Nakazato H. and Edmonds M. 1972. The isolation and purification of rapidly labeled polysome-bound ribonucleic acid on polythymidylate cellulose. *J. Biol. Chem.* **247:** 3365–3367.

Nichols B.L., Dudley M.A., Nichols V.N., Putman M., Avery S.E., Fraley J.K., Quaroni A., Shiner M., and Carrazza F.R. 1997. Effects of malnutrition on expression and activity of lactase in children. *Gastroenterology* **112:** 742–751.

Novoradovskaya N., Payette T., Novoradovsky A., Braman J., Chin N., Pergamenschikov A., Fero M., and Botstein D. 2000. Pooled, high quality reference RNA for human microarrays. *Strategies Newsletter* **13:** 121–122. Stratagene, Inc., LaJolla, California.

Okamuro J.K. and Goldberg R.B. 1989. Regulation of plant gene expression: General principles. In *The biochemistry of plants: A comprehensive treatise* (ed. A. Marcus et al.), vol. 15, pp.1–82, Academic Press, New York.

Ohnmacht G.A., Wang E., Mocellin S., Abati A., Fetsch P.A., Rosenberg S.A., and Marincola F.M. 2001. Short-term kinetics of tumor antigen expression *in vivo*. *J. Immunol.* **167:** 1809–1820.

Panchuk-Voloshina N., Haugland R.P., Bishop-Stewart J., Bhalgat M.K., Millard P.J., Mao F., and Leung W.Y. 1999. Alexa dyes, a series of new fluorescent dyes that yield exceptionally bright, photostable conjugates. *J. Histochem. Cytochem.* **47:** 1179–1188.

Pease A.C., Solas D., Sullivan E.J., Cronin M.T., Holmes C.P., and Fodor S.P. 1994. Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proc. Natl. Acad. Sci.* **91:** 5022–5026.

Penman S., Fan H., Perlman S., Rosbach M., Weinberg R., and Zylber E. 1970. Distinct RNA synthesis systems of the HeLa cells. *Cold Spring Harbor Symp. Quant. Biol.* **35:** 561–575.

Perou C.M., Jeffrey S.S., van de Rijn M., Rees C.A., Eisen M.B., Ross D.T., Pergamenschikov A., Williams C.F., Zhu S.X., Lee J.C., Lashkari D., Shalon D., Brown P.O., and Botstein D. 1999. Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proc. Natl. Acad. Sci.* **96:** 9212–9217.

Perou C.M., Sorlie T., Eisen M.B., van de Rijn M., Jeffrey S.S., Rees C.A., Pollack J.R., Ross D.T., Johnsen H., Akslen L.A., Fluge O., Pergamenschikov A., Williams C., Zhu S.X., Lonning P.E., Borresen-Dale A.L., Brown P.O., and Botstein D. 2000. Molecular portraits of human breast tumours. *Nature* **406:** 747–752.

Perry R.P. and Kelley D.E. 1972. The production of ribosomal RNA from high molecular weight precursors. 3. Hydrolysis of pre-ribosomal and ribosomal RNA by a 3′-OH specific exoribonuclease. *J. Mol. Biol.* **70:** 265–279.

Phillips J. and Eberwine J.H. 1996. Antisense RNA amplification: A linear amplification method for analyzing the mRNA population from single living cells. *Methods* **10:** 283–288.

Pichersky E., Bernatzky R., Tanksley S.D., and Cashmore A.R. 1986. Evidence for selection as a mechanism in the concerted evolution of *Lycopersicon esculentum* (tomato) genes encoding the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase. *Proc. Natl. Acad. Sci.* **83:** 3880–3884.

Prasad V.R. 1993. Genetic analysis of retroviral reverse transcriptase structure and function. In *Reverse transcriptase* (ed. A.M. Skalka and S.P. Goff), pp. 135–162. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Rajeevan M.S., Dimulescu I.M., Unger E.R., and Vernon S.D. 1999. Chemiluminescent analysis of gene expression on high-density filter arrays. *J. Histochem. Cytochem.* **47:** 337–342.

Randolph J.B. and Waggoner A.S. 1997. Stability, specificity and fluorescence brightness of multiply-labeled fluorescent DNA probes. *Nucleic Acids Res.* **25:** 2923–2929.

Reymond P., Weber H., Damond M., and Farmer E.E. 2000. Differential gene expression in response to mechanical wounding and insect feeding in *Arabidopsis. Plant Cell* **12:** 707–719.

Richmond T. and Somerville S. 2000. Chasing the dream: Plant EST microarrays. *Curr. Opin. Plant Biol.* **3:** 108–116.

Richmond C.S., Glasner J.D., Mau R., Jin H., and Blattner F.R. 1999. Genome-wide expression profiling in *Escherichia coli* K-12. *Nucleic Acids Res.* **27:** 3821–3835.

Riley M. and Labedan B. 1996. *Escherichia coli* gene products: Physiological functions and common ancestries. In Escherichia coli *and* Salmonella*: Cellular and molecular biology*, 2nd edition (ed. F.C. Neidhardt et al.), pp. 2118–2202. American Society for Microbiology, Washington, D.C.

Roozemond R.C. 1976. Ultramicrochemical determination of nucleic acids in individual cells using the Zeiss UMSP-I microspectrophotometer. Application to isolated rat hepatocytes of different ploidy classes. *Histochem. J.* **8:** 625–638.

Ross D.T., Scherf U., Eisen M.B., Perou C.M., Rees C., Spellman P., Iyer V., Jeffrey S.S., Van de Rijn M., Waltham M., Pergamenschikov A., Lee J.C., Lashkari D., Shalon D., Myers T.G., Weinstein J.N., Botstein D., and Brown P.O. 2000. Systematic variation in gene expression patterns in human cancer cell lines. *Nat. Genet.* **24:** 227–235.

Ruan Y., Gilmore J., and Conner T. 1998. Towards *Arabidopsis* genome analysis: Monitoring expression profiles of 1400 genes using cDNA microarrays. *Plant J.* **15:** 821–833.

Sambrook J. and Russell D.W. 2001. *Molecular cloning: A laboratory manual*, 3rd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Schena M. and Davis R.W. 2000. Technology standards for microarray research. In *Microarray biochip technology* (ed. M. Schena), pp. 1–18. Eaton Publishing/ BioTechniques Book Division, Natick, Massachusetts.

Schena M., Shalon D., Davis R.W., and Brown P.O. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270:** 467–470.

Schena M., Shalon D., Heller R., Chai A., Brown P.O., and Davis R.W. 1996. Parallel human genome analysis: Microarray-based expression monitoring of 1000 genes. *Proc. Natl. Acad. Sci.* **93:** 10614–10619.

Schick B.P. and Eras J. 1995. Proteoglycans partially co-purify with RNA in TRI reagent and can be transferred to filters by northern blotting. *BioTechniques* **18:** 574–578.

Schneiderbauer A., Sandermann Jr. H., and Ernst D. 1991. Isolation of functional RNA from plant tissues rich in phenolic compounds. *Anal. Biochem.* **197:** 91–95.

Selinger D.W., Cheung K.J., Mei R., Johansson E.M., Richmond C.S., Blattner F.R., Lockhart D.J., and Church G.M. 2000. RNA expression analysis using a 30 base pair resolution *Escherichia coli* genome array. *Nat. Biotechnol.* **18:** 1262–1268.

Sevag M.G., Lackman D.B., and Smolens J. 1938. The isolation of the components of streptococcal nucleoproteins in serologically active form. *J. Biol. Chem.* **1254:** 425–436.

Sherman F. 1991. Getting started with yeast. *Methods Enzymol.* **194:** 3–21.

Smulski D.R., Huang L.L., McCluskey M.P., Reeve M.J., Vollmer A.C., Van Dyk T.K., and LaRossa R.A. 2001. Combined, functional genomic-biochemical approach to intermediary metabolism: Interaction of acivicin, a glutamine amidotransferase inhibitor, with *Escherichia coli* K-12. *J. Bacteriol.* **183:** 3353–3364.

Spremuli L. 2000. Protein synthesis, assembly and degradation. In *Biochemistry and molecular biology of plants* (ed. B.B. Buchanan et al.), pp. 412–454, American Society of Plant Physiologists, Rockville, Maryland.

Stinson S.F., Alley M.C., Kopp W.C., Fiebig H.H., Mullendore L.A., Pittman A.F., Kenney S., Keller J., and Boyd M.R. 1992. Morphological and immunocytochemical characteristics of human tumor cell lines for use in a disease-oriented anticancer drug screen. *Anticancer Res.* **12:** 1035–1053.

Stoltenburg R., Wartmann T., Kunze I., and Kunze G. 1995. Reliable method to prepare RNA from free and membrane-bound polysomes from different yeast species. *BioTechniques* **18:** 564–568.

Stulnig T.M. and Amberger A. 1994. Exposing contaminating phenol in nucleic acid preparations. *BioTechniques* **16:** 402–404.

287

Suganuma A. and Gupta K.C. 1995. An evaluation of primer length on random-primed DNA synthesis for nucleic acid hybridization: Longer is not better. *Anal. Biochem.* **224:** 605–608.

Sugiura M. and Takeda Y. 2000. Nucleic acids. In *Biochemistry and molecular biology of plants* (ed. B.B. Buchanan et al.), pp. 260–310. American Society of Plant Physiologists, Rockville, Maryland.

Takahashi N. and Ko M.S. 1994. Toward a whole cDNA catalog: Construction of an equalized cDNA library from mouse embryos. *Genomics* **23:** 202–210.

Talaat A.M., Hunter P., and Johnston S.A. 2000. Genome-directed primers for selective labeling of bacterial transcripts for DNA microarray analysis. *Nat. Biotechnol.* **18:** 679–682.

Tanese N. and Goff S.P. 1988. Domain structure of the Moloney murine leukemia virus reverse transcriptase: Mutational analysis and separate expression of the DNA polymerase and RNase H activities. *Proc. Natl. Acad. Sci.* **85:** 1777–1781.

Tanese N., Roth M., and Goff S.P. 1985. Expression of enzymatically active reverse transcriptase in *Escherichia coli*. *Proc. Natl. Acad. Sci.* **82:** 4944–4948.

Taniguchi M., Miura K., Iwao H., and Yamanaka S. 2001. Quantitative assessment of DNA microarrays—Comparison with northern blot analyses. *Genomics* **71:** 34–39.

Tavangar K., Hoffman A.R., and Kraemer F.B. 1990. A micromethod for the isolation of total RNA from adipose tissue. *Anal. Biochem.* **186:** 60–63.

Trenkle T., Welsh J., Jung B., Mathieu-Daude F., and McClelland M. 1998. Non-stoichiometric reduced complexity probes for cDNA arrays. *Nucleic Acids Res.* **26:** 3883–3891.

Uemura E. 1980. Age-related changes in neuronal RNA content in rhesus monkeys (*Macaca mulatta*). *Brain Res. Bull.* **5:** 117–119.

Ullrich A., Shine J., Chirgwin J., Pictet R., Tischer E., Rutter W., and Goodman H. 1977. Rat insulin genes: Construction of plasmids containing the coding sequences. *Science* **196:** 1313–1319.

van Gelder R.N., von Zastrow M.E., Yool A., Dement W.C., Barchas J.D., and Eberwine J.H. 1990. Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proc. Natl. Acad. Sci.* **87:** 1663–1667.

Wang Y., Rea T., Bian J., Gray S., and Sun Y. 1999. Identification of the genes responsive to etoposide-induced apoptosis: Application of DNA chip technology. *FEBS Lett.* **445:** 269–273.

Wang E., Miller L.D., Ohnmacht G.A., Liu E.T., and Marincola F.M. 2000. High-fidelity mRNA amplification for gene profiling. *Nat. Biotechnol.* **18:** 457–459.

Warburg O. and Christian W. 1942. Isolierung und kristallisation des gärungsferments enolase. *Biochem. Z.* **310:** 384-421.

Wei Y., Lee J.M., Richmond C., Blattner F.R., Rafalski J.A., and LaRossa R.A. 2001. High-density microarray-mediated gene expression profiling of *Escherichia coli*. *J. Bacteriol.* **183:** 545–556.

Welford S.M., Gregg J., Chen E., Garrison D., Sorensen P.H., Denny C.T., and Nelson S.F.1998. Detection of differentially expressed genes in primary tumor tissues using representational differences analysis coupled to microarray hybridization. *Nucleic Acids Res.* **26:** 3059–3065.

Wendisch V.F., Zimmer D.P., Khodursky A., Peter B., Cozzarelli N., and Kustu S. 2001. Isolation of *Escherichia coli* mRNA and comparison of expression using mRNA and total RNA on DNA microarrays. *Anal. Biochem.* **290:** 205–213.

Wessendorf M.W. and Brelje T.C. 1992. Which fluorophore is brightest? A comparison of the staining obtained using fluorescein, tetramethylrhodamine, lissamine rhodamine, Texas red, and cyanine 3.18. *Histochemistry* **98:** 81–85.

Wildsmith S.E., Archer G.E., Winkley A.J., Lane P.W., and Bugelski P.J. 2001. Maximization of signal derived from cDNA microarrays. *BioTechniques* **30:** 202–208.

Wilfinger W.W., Mackey K., and Chomczynski P. 1997. Effect of pH and ionic strength on the spectrophotometric assessment of nucleic acid purity. *BioTechniques* **22:** 474–476, 478–481.

Williamson R., Morrison M., Lanyon G., Eason R., and Paul J. 1971. Properties of mouse globin messenger ribonucleic acid and its preparation in milligram quantities. *Biochemistry* **10:** 3014–3021.

Yang G.P., Ross D.T., Kuang W.W., Brown P.O., and Weigel R.J. 1999. Combining SSH and cDNA microarrays for rapid identification of differentially expressed genes. *Nucleic Acids Res.* **27:** 1517–1523.

Yu H., Chao J., Patek D., Mujumdar R., Mujumdar S., and Waggoner A.S. 1994. Cyanine dye dUTP analogs for enzymatic labeling of DNA probes. *Nucleic Acids Res.* **22:** 3226–3232.

Yue H., Eastman P.S., Wang B.B., Minor J., Doctolero M.H., Nuttall R.L., Stack R., Becker J.W., Montgomery J.R., Vainer M., and Johnston R. 2001. An evaluation of the performance of cDNA microarrays for detecting changes in global mRNA expression. *Nucleic Acids Res.* **29:** E41.

Zejie Y., Ziaoyi W., Yu T., and Huan H. 1998. The method of microdisplacement measurement to improve the space resolution of array detector. *Med. Eng. Phys.* **20:** 149–151.

Zhu T. and Wang X. 2000. Large-scale profiling of the *Arabidopsis* transcriptome. *Plant Physiol.* **124:** 1472–1476.

Zhu Z. and Waggoner A.S. 1997. Molecular mechanism controlling the incorporation of fluorescent nucleotides into DNA by PCR. *Cytometry* **28:** 206–211.

Zhu Z., Chao J., Yu H., and Waggoner A.S. 1994. Directly labeled DNA probes using fluorescent nucleotides with different length linkers. *Nucleic Acids Res.* **22:** 3418–3422.

因特网资源

288

http://afgc.stanford.edu Arabidopsis Functional Genomics Consortium (AFGC)

http://cmgm.stanford.edu/pbrown/mguide/index.html The MGuide Version 2.0. The Brown Lab's complete guide to microarraying for the molecular biologist.

http://www.ambion.com Ambion, Inc. homepage

http://www.apbiotech.com Amersham Biosciences homepage

http://www.atcc.org American Type Culture Collection homepage

http://www.clontech.com BD Biosciences Clontech homepage

http://www.lifetech.com Invitrogen homepage

http://www.microarrays.org/protocols.html Microarray protocols maintained at the University of California, San Francisco, Department of Biochemistry and Biophysics.

http://www.millipore.com Millipore Corporation homepage

http://www.molecularprobes.com Molecular Probes, Inc. homepage

http://www.mrcgene.com Molecular Research Center homepage

http://www.qiagen.com QIAGEN homepage

第4章　膜点样cDNA阵列

289

Kevin G. Becker，William H. Wood III，Chris Cheadle
（DNA阵列研究部，国家衰老研究所，国立卫生院，巴尔的摩，马里兰州，21224）

导言

现代微阵列的许多概念和策略在20世纪80年代得到了发展，当时滤膜被用在使用复杂的靶分子对互补DNA（cDNA）文库进行差异筛选中（Sargent and Dawid 1983；Friemert et al. 1989；Gress et al. 1992）。在得到了经过测序的cDNA文库以后，利用基于滤膜的方法对来自于人类胎儿大脑、*Drosophila*和鼠胚胎以及其他组织的表达谱进行了分析（Lennon and Lehrach 1991；Gress et al. 1992；Maier et al. 1994；Takahashi and Ko 1994）。不久以后，自动点样的膜阵列（Maier et al. 1994）、玻片cDNA微阵列（Schena et al. 1995）和寡核苷酸"芯片"（Lockhart et al. 1996）的发展，使得能够对基因表达进行更为有效的系统比较。

290

膜阵列已经成功地得到了商品化，这主要是因为膜技术容易整合到现有的实验室方案中去（见 www. resgen. com；www. clontech. com）。对于普通的分子生物学实验室，商业化生产的膜阵列既提供了进行基因组规模基因表达研究的"直通车"，又不需要购买大量的cDNA克隆或者专用的点样和扫描仪器。除此之外，定制的膜阵列提供了一种灵活的形式，可以自制并用作商业化产品的经济型替代。

玻璃和膜点样微阵列，两者都可以在实验室内制备，有时被看为两种互为替代的技术。实际上，每种形式都有自己的长处和不足，两种技术应该被看成是互补的，而非竞争性的可选方法。基因阵列在玻片（微阵列）上的点样密度要比在膜（大型的阵列）上高。

更重要的是，同时进行双色杂交反应时，在玻片微阵列上能够使用内部的参考探针作为制备微阵列时芯片与芯片之间差异的对照。能够包含该类型的对照是这两种系统的一个主要区别。在应用上，玻片和膜阵列的相对灵敏度可能取决于阵列的质量和研究者制备在杂交时产生较低背景的标记靶分子的能力。膜点样/使用放射性标记检测的阵列的潜力包括：

- **放射性前体标记的 cDNA 比荧光化合物标记具有更高的特异活性**。放射性标记 dNTP 的掺入效率比 Cy5-或 Cy3-标记 dNTP 高，这使得标记时所用的 RNA 量要少得多。用荧光 dNTP 进行标记时一般需要 50～200 μg 的总 RNA，而使用 [^{33}P]dNTP 标记时只需要 1～10 μg。从这个意义上讲，放射性标记的靶分子与膜阵列杂交的灵敏度要比荧光标记的靶分子与玻片微阵列杂交的灵敏度高出 100 倍（Bertucci et al. 1999）。尽管不能将一个实验所需要的 RNA 的量和检测基因表达差异的能力相混淆，然而，膜阵列可以在不扩增靶分子的情形下对实验进行重复。增加重复杂交的次数能够改善实验的统计学显著性。（Andrews et al. 2000；Barrett et al. 2001）（关于进一步的讨论，见第 7 章，PART Ⅰ）
- **应用单一标记系统（^{33}P）比双重荧光标记具有优势**。Cy3-和 Cy5-标记的 dUTP 掺入 DNA 的反应速率各不相同，并且 Cy5-标记的 dUTP 掺入效率较低。掺入效率的差异会带来需要通过标准化算法进行校正的偏差，这一处理过程可能会引入误差。不仅如此，对一些基因进行的差异表达检测表现出了对所应用染料组合具有敏感性（Taniguchi et al. 2001）。
- **尼龙膜阵列的制备和处理相对快捷和简便**。比较而言，玻片阵列的制备，包括载玻片的包被和老化、点样、变性以及化学处理，是一个冗长的过程。某些制备膜阵列的阵列设备中可以使用未经纯化的聚合酶链反应（PCR）产物，从而极大地缩短了探针集制备的工作量。将一组大约 1000 个浓缩的 PCR 产物点样到尼龙膜上，一般需要一天时间以进行膜的准备、点样和 UV 交联。
- 291 **许多原因使得膜阵列的成本相当地低**，包括能够使用未经纯化的 PCR 产物进行点样，使用常见的标记和杂交实验方案，放射性标记的核苷酸与荧光标记核苷酸相比花费相对较少，膜阵列可以重复杂交 3～5 次，使用标准的放射性检测方法（胶片或者磷像屏）。总而言之，这些因素使得系统建立和使用的成本较低，并且缩短了从系统建立到产出数据的时间。

作为总结，基于膜 cDNA 阵列的基因表达研究已经被证实具有高灵敏、可重复、成本低并且灵活性高等特点。与其他阵列形式相比，它们一般只要求专业程度较低且不很昂贵的设备和试剂。根据作者应用 cDNA 阵列关键设施的经验，尼龙膜对实验操作者培训的需求较少，并且容易从专门的阵列设施转移到标准的分子学实验室或者是合作实验室中。这些优势使得可以进行人员的快速培训，并花费较少的时间来产出高质量的微阵列数据。

292 PART Ⅰ 膜阵列的制备和杂交

本章的第一部分介绍了使用基于膜的阵列建立微阵列实验的一系列步骤。首先，将根据第 1 章中介绍方法进行分离的扩增后的 cDNA 克隆，用标准的自动点样仪点样在

尼龙膜上，再通过 UV 照射固定到膜上。通常情况下，这里的点样过程要比点样到固相支持物表面（在第 2 章中介绍）更为直接。作为替代，可以从 ResGen（www.resgen.com）或者 CLONTECH（www.clontech.com）购买到能够点样于可重复使用尼龙膜上的哺乳动物和酵母克隆集，或者是可读框（ORF）。然后，将根据第 3 章 PART Ⅰ中介绍的方法所制备的靶分子 RNA，利用反转录酶直接掺入 [α-^{33}P]dCTP 进行放射标记。在模板 RNA 被水解成碱基后，使用离心柱色谱法纯化标记的 cDNA 靶分子。然后将标记的 DNA 靶分子与膜微阵列在与传统的 Southern 杂交相似的条件下进行杂交，通常在高严格度条件下进行。杂交结果使用标准放射性同位素检测（磷像）技术成像，并按照 PART Ⅱ中介绍的方法进行分析。

下面一系列方案介绍了：

- 将 cDNA 阵列点样到尼龙膜上和后续的处理（方案 1）。
- RNA 的制备和放射性标记以生成靶分子（方案 2）。
- 标记靶分子和膜阵列的杂交（方案 3）。

方案 1

293

膜阵列的点样

使用自动点样仪将 cDNA 阵列点样到尼龙膜上，在某种程度上要比点样到玻片表面上更为宽松。由于核酸与尼龙膜之间能够高效地结合，而且使用的实验方案是经过多年应用发展成熟的，所以点样优化所需要的时间较少。在膜上点样时最好使用钝头的点样针，而不是分叉的或者鹅毛笔针类型的点样针，它们会损伤滤膜。此外，鹅毛笔针在滤膜表面的毛细作用会产生各种大小的样点，而钝头的点样针可以产生直径和 DNA 含量均一的样点。实心针的另一个优点是，可以直接用通常会堵塞鹅毛笔针的未纯化的 PCR 产物来点样。在作者的实验室里，使用的是具有针和环点样系统的 Affymetrix 417 点样仪。该系统单次蘸取 PCR 产物后可以在超过 42 个阵列上点样产生均一的样点。有关点样系统更详细的情况，特别是针和环的设定，见第 2 章中**用于微阵列接触点样的点样仪和点样针**信息栏。

滤膜必须放置在一个平稳的位置以进行点样。在作者的实验室里，制作了与标准显微镜载玻片（75 mm×175 mm×1.5 mm）同样大小的不锈钢模具，将膜临时固定在这些模具上进行点样。Affymetrix 417 点样仪的改装以及在点样时使用不锈钢模具代替显微镜载玻片进行膜的固定见图 4-1。

膜阵列探针的制备：cDNA 克隆的 PCR 扩增

可以用第 1 章（用于点样阵列的探针）中介绍的方法制备在膜上点样的材料。除了不需要对 PCR 扩增产物进行纯化之外，cDNA 克隆的扩增按照第 1 章，方案 2 的介绍进行操作。标准的 PCR 扩增完成后，100 μl 的反应蒸发浓缩至大约 24 μl，这可以通过将反应再次放入 96 孔 PCR 仪中，并将盖子打开后在 70℃下加热 1～2 h 来实现。

蒸发与PCR产物的乙醇沉淀或柱纯化不同，不需要对扩增子进行额外转移或处理。与玻璃基底不同，尼龙膜容许点样混合物中有微量污染物杂质存在，如游离的核苷酸、酶和残留的PCR引物。然而，未纯化PCR扩增产物只能用于那些可以容许不纯物质而不会堵塞点样针的阵列系统，如上面叙述的基于针和环系统的仪器。要按照第1章，方案4中所述，对PCR产物的产量及单一条带的扩增进行检查。

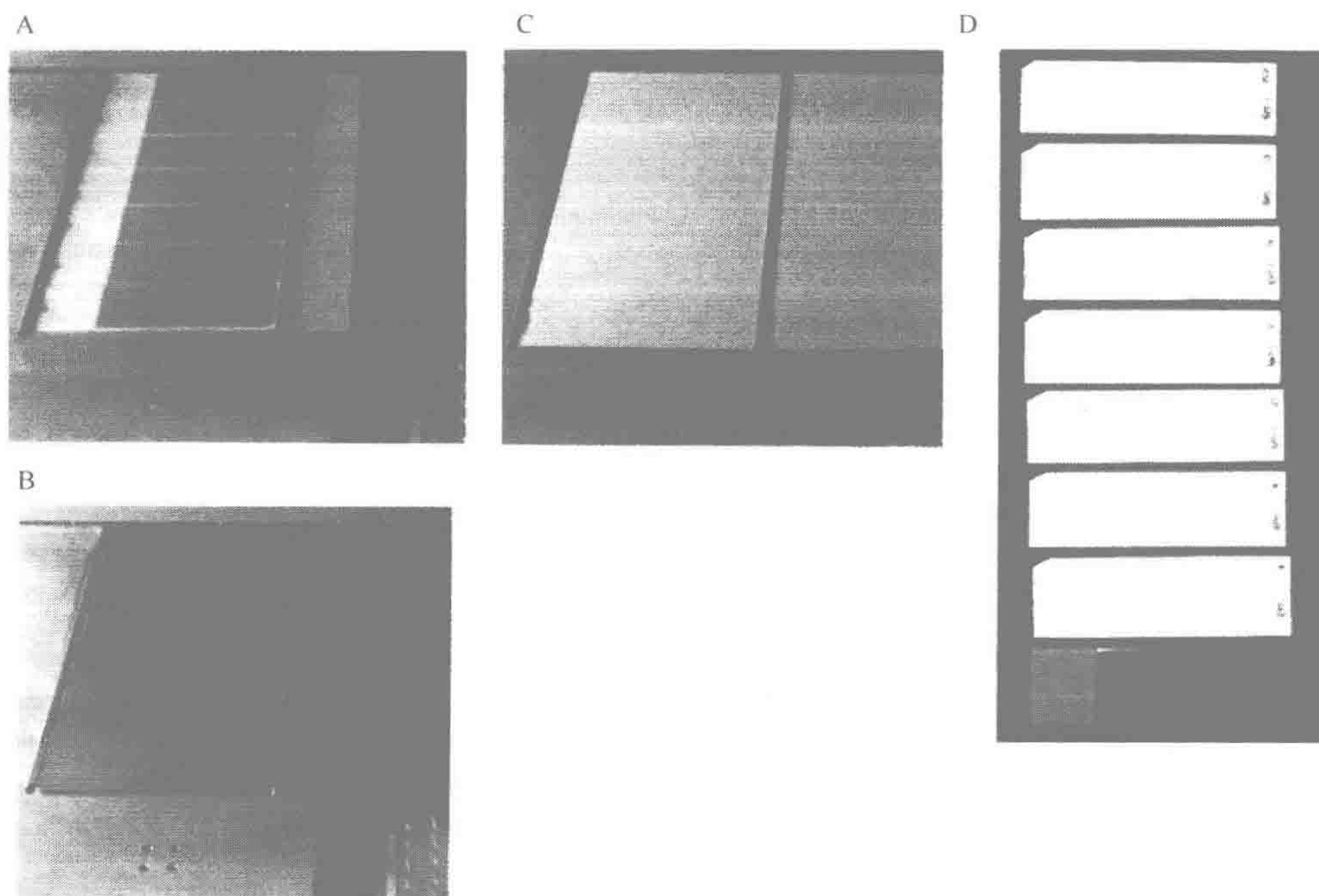


图 4-1 改装 Affymetrix 417 点样仪以适合在尼龙膜上点样

A. 加载了7片标准显微镜载玻片的阵列点样平台的局部；B. 被不锈钢模具取代的同一个7片载玻片的区域；C. 黏有尼龙膜的不锈钢模具；D. 7片在点样后切割和标记的75 mm×25 mm尼龙膜阵列

材料

注意：标记有〈!〉的材料的正确处理方法参见附录3。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录2。
使用前将储存液稀释到合适的浓度。
NaOH(1 mol/L)〈!〉

核酸和寡核苷酸

cDNA 探针

按照第1章，方案2中介绍的方法制备探针。关于制备和浓缩探针用于膜阵列点样的更多信息，见本方案开始部分的信息栏。

专用设备

点样仪（如 Affymetrix 417）
Koh-I-Noor rapidograph 记号笔（使用 Rapid-draw 3084-F 墨水）
Krylon 不干胶（7020，Krylon，Inc.）
无粉丁腈手套
Nytran+ Supercharge 尼龙膜（10416292，Schleicher & Schuel）
塑料多孔板密封膜
放置尼龙膜用于点样的不锈钢模具（75 mm×175 mm×1.5 mm）

这些模具可以在机械商店里按要求规格进行定制

Stratalinker 2400 交联仪（Stratagene）

方法

膜的制备

1. 戴上无粉的丁腈手套，将 Nytran+ Supercharge 尼龙膜切割成适合于每 6 个 75 mm×175 mm×1.5 mm不锈钢模具的大小。

   最好切割带有直边的尼龙膜，并且刀片穿过膜的垫纸进行切割。

2. 将 Krylon 不干胶轻轻地喷洒在每个不锈钢模具的表面，这种黏合剂保持黏性并且不会变干。给每个模具放置上尼龙膜，并轻压膜的垫纸使膜紧贴在模具上。
3. 将模具/膜放置在点样仪中合适的平台位置上，覆盖全部的点样区域。确保每件模具/膜单元与平台之间保持平整，并且在整个平台上没有任何高出的点。

在尼龙膜上进行阵列点样

4. 在开始点样每组三块 96 孔微量滴定板之前，向每份 cDNA 样品（通常约 20 μl）中加入 1 mol/L NaOH（约 5 μl），使得最终浓度为 0.2 mol/L。轻轻涡旋多孔板数秒钟以混匀样品。短暂离心收集样品。

   点过一次样的 96 孔板中不需要额外的 NaOH，但是可能需要加入 5～10 μl 的水以便对蒸发进行校正。孔内的材料最少为 10 μl，以确保样品蘸取时环可以充满样品。

5. 将最初的三块多孔板放入点样仪中。
6. 使用 300 μm 的点样针进行探针的点样，各点之间的中心距离为 665 μm。

   这一过程可以在相当于 25 mm×75 mm 的显微镜载玻片大小的区域上点样 2304 个独立的点，并且从 24 块 96 孔板可以产生 16 个一致的 12×12 方格或者子阵列。在许多情况下，作者在同样的区域上点样重复的两份 12 块多孔板，这样可以点样带有重复的 1152 个点。制备 42 个同样的阵列需要大约 8 h 的连续点样。关于阵列点样更多的详细信息，见第 2 章，方案 2。

296

7. 每组三块多孔板点样结束后，从点样仪中取出 96 孔板，用塑料多孔板密封膜密封，并将其在－20 ℃下保存。

   一次标准的 PCR 可用来点样 5～10 次，总计得到约 200～400 个单独的阵列。在理想情况下，应该以所有点样用材料保持一致的方式来进行点样，以减少与长期储存有关的假

象。PCR 板在－20 ℃下数月的长期储存似乎会导致衰弱的信号。尚不清楚为何会出现这一现象；然而，这看上去并不是因为 DNA 保存在 NaOH 中而造成的。

膜的点样后处理

点样后处理需要大约 1 h，并得到可以立刻用于杂交的阵列。

8. 在点样操作完成后，将仍旧黏附了已点样膜的不锈钢模具放到 Strataliker 2400 UV 交联仪中，在 120 mJ/s 照射下，UV 交联约 1 min，使得已经点样的 DNA 固定到膜上。
9. 用剃须刀片分开黏附了膜的不锈钢模具。
10. 以使用 Rapid-draw 3084-F 墨水的 Koh-I-Noor rapidograph 记号笔标记膜，墨水中结合了放射标记的靶，因此，可以从扫描图像中对滤膜进行识别。

如果点样后的膜不立即使用，将其夹在标准的复印纸中在室温下保存。

297

方案 2

RNA 的制备和标记

关于分离总 RNA 用于标记的方法，见第 3 章 PART Ⅰ。使用反转录酶（RT）直接掺入［α-^{33}P］dCTP 进行 cDNA 靶分子的制备（Whitney et al. 1999）。下面的方案看来对抑制荧光标记的微量污染物不敏感，也许是因为两种放射性标记核苷酸的掺入效率要更高一些。在一个典型反应中，每份样品使用 5 μg 的总 RNA；如果使用长时间的曝光，从少至 1 μg 的总 RNA 也可以获得有用信号，而使用超过 10 μg 的 RNA 并不会产生更多的信息。

如果使用的膜比这里叙述的尺寸要大，则有必要按照增加的杂交液的用量来相应增加 RNA 的用量。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。
使用前将储存液稀释到合适的浓度。

二硫苏糖醇（DTT）（0.1 mol/L）

EDTA(0.5 mol/L，pH 8.0)

NaOH(0.1 mol/L)〈!〉

Tris(1 mol/L，pH 8.0)

酶和缓冲液

5×第一链缓冲液（18084-014，GIBCO）

RNaseOUT(10777-019，GIBCO)
SUPERSCRIPT Ⅱ RT(18084-014，GIBCO)

核酸和寡核苷酸

[α-^{33}P]dCTP(3000 Ci/mmol)58430，ICN〈!〉
dNTP 溶液，不含有 dCTP，每种浓度均为 8 mmol/L(27-20X0-0，Pharmacia)
Oligo$(dT)_{10-20}$引物（1 μg/μl）(Poly TGF，Research Genetics)
总 RNA 制备物

总 RNA 制备按照第 3 章，PART Ⅰ介绍的方案之一制备 RNA 制备物。

专用设备

Bio-Spin P-30 离心柱（Bio-Rad）

方法

298

cDNA 第一链的合成

1. 将 10 μg 的总 RNA 与 1 μl 的 oligo$(dT)_{10-20}$引物（1 μg/μl）混合，并加入 H_2O 至体积为 15 μl 以建立复性反应。将混合物在 65℃下加热 5 min，然后在冰上孵育 2 min。
2. 向15 μl 的复性反应中加入下列成分，准备反转录反应混合物：

5×第一链缓冲液	8 μl
20 mmol/L dNTP 溶液（不含有 dCTP）	4 μl
0.1 mol/L DTT	4 μl
RNaseOUT(40 单位)	1 μl
[α-^{33}P]dCTP(3000 Ci/mmol)	5 μl
H_2O	至终体积为 38 μl

3. 向反应中加入 2 μl SUPERSCRIPT Ⅱ RT(200 单位/μl)，并将混合物在 42 ℃下孵育 35min。

   反转录酶对在气/液界面的变性非常敏感，因此当混合反应的组分时，要非常小心避免产生泡沫。有时这很难做到，因为许多反转录酶制备物的缓冲液中含有非离子型去污剂，如 Nonidet P-40 或者 Triton X-100。

4. 再次加入 2 μl SUPERSCRIPT Ⅱ RT(200 单位/μl)，并将混合物在 42 ℃下孵育 35 min。

RNA 的水解

5. 加入 5 μl 的 0.5mol/L EDTA 以终止反应。

   加入 EDTA 之后，在继续加入 NaOH（步骤 6）之前，确保反应混合物混匀。核酸会在碱性镁溶液中发生沉淀。

6. 加入 10 μl 0.1mol/L NaOH，充分混匀并短暂离心，然后将样品放在 65℃下孵

育 30 min，以使模板 RNA 被水解。

7. 加入 25 μl 1mol/L Tris（pH 8.0）以中和样品。

使用 Bio-Rad Bio-Spin P-30 离心柱净化标记的靶分子

8. 将柱子放入微量离心管中短暂离心使树脂铺平。
9. 去掉离心柱帽并刺破离心柱底部。
10. 将离心柱放入 2 ml 的收集管中。
11. 室温下在微型离心机中，将柱子低速（1000 r/min）离心 2 min。
12. 弃去流出液体并重复离心（步骤 11）直到除去柱子中的所有液体。
13. 将离心柱放入一支新的微量离心管内。
14. 将标记的样品（不超过 82 μl）加入到柱子上。

299

15. 在微型离心机中，将柱子低速（1000 r/min）离心 4 min，并回收流出液体（约 82 μl）。低速离心（1000 r/min）4 min 并再次上样（约 82 μl）

    柱子可以在−20 ℃下保存。

16. 将 1 μl 的小份样品在闪烁计数器中计数。

    一次典型的标记反应（75 μl）通常得到 3×10^7～5×10^7 总 cpm counts per minute（注：液体闪烁计数时，每分钟的计数。由于液闪仪效率不同，建议采用每分钟衰变数 dpm 为单位）。

300

方案 3

靶分子与膜阵列的杂交

一般情况下，膜微阵列的杂交在与传统的 Southern 杂交相似的条件下进行。使用高离子强度的缓冲液以促进互补链的结合，同时使用各种封闭试剂来降低背景干扰。杂交和清洗一般在高严格度的条件下进行，以抑制交叉杂交。关于杂交反应的更多详细信息和特别注意事项，请参考第 3 章的导言部分和方案 20。在本方案中，使用［α-^{33}P］dCTP 进行放射性标记的 cDNA 靶分子（方案 2）与点样的膜阵列进行杂交。杂交（和后续分析）结束后，可以将靶分子从膜上剥离下来，膜能够重新应用于后续的杂交反应，如同本方案结束部分的附加方案中介绍的那样。

各种问题以及可能的解决方法在图 4-3 中的疑难问题解答指南中给出。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

氯仿〈!〉

DEPC 处理的 H_2O〈!〉(见**抑制 RNase** 的信息栏)
硫酸葡聚糖（50%）
2×SSC
1×SSC/0.1% SDS〈!〉
2×SSC/0.1% SDS
Microhyb(Hyb250.GF，Research Genetics)

核酸和寡核苷酸

DNA 阵列，按照方案 1 的介绍点样到尼龙膜上
人 C_0t-1 DNA(1 mg/ml)（15279-001，LTI，Inc.）
poly(dA)（8 mg/ml)（P 9403，Sigma)
cDNA 靶分子（约 75 μl，3×10^7～5×10^7 cpm)，按照方案 2 的介绍进行标记

专用设备

Falcon 试管（35-2098，Falcon)
玻璃或金属平板
见步骤 8（译者注：原书印刷有误，应为步骤 10）
旋转杂交炉（如 HS9360，HybAid，Inc.）
磷像仪（如 Molecular Dynamics STORM)
磷像屏（如 Molecular Dynamics)

方法

膜的预清洗

1. 将点样在 25 mm×75 mm 尼龙膜上的 cDNA 阵列放入容量为 50 ml 的空的一次性 Falcon 试管中，并将膜上点过样的一面朝向试管内部。

   最多可以环绕管内壁放置 3 块膜，而不至于在与靶分子杂交时发生膜之间的相互重叠。

2. 向试管中加入 50 ml 的 2×SSC 进行滤膜的预洗。弃去液体，并将试管倒置于纸巾上，沥干。

   该步骤使得膜附着在试管内壁上。

3. 将阵列的边缘对齐并排列好，用 5 ml 塑料吸管紧贴着膜滚动以除去膜下方的所有气泡。

阵列杂交反应

4. 混合下列各成分制备杂交缓冲液：

含有 10%硫酸葡聚糖（最终浓度）的 Microhyb	4 ml
1 mg/ml 人 C_0t-1 DNA，使用前在 95 ℃下变性处理 5 min	10 μl
8 mg/ml poly(dA) 使用前在 95 ℃下变性处理 5 min	10 μl

5. 将杂交液加入到装有阵列的Falcon试管中，在标准的旋转杂交炉中于42 ℃将膜预杂交2 h。
6. 预杂交后，将全部体积的标记探针（约75 μl，3×10^7～5×10^7 cpm）在95 ℃下变性处理5 min，然后加入装有膜阵列和杂交缓冲液（约4 ml）的Falcon试管中（步骤5）。
7. 将试管在42 ℃下旋转孵育12～18 h，进行杂交反应。
8. 杂交结束后，将杂交液倒入一适于在－20 ℃下保存的管中储藏。

   尽管靶分子看上去似乎可以被有效地重复利用，但是直到本文撰写之时，对重复利用的杂交液的使用效率还没有进行广泛的研究。因此，研究者应该自行确定重复利用的有效性。按照上述方法进行膜的预杂交。将储存的杂交液在100 ℃下变性处理5 min，冷却至42 ℃并保持5 min，然后将溶液加入到已经沥去杂交缓冲液的滤膜上。

膜的清洗

重要事项：在清洗过程的任何时刻，不要使膜干燥。建议膜清洗时的温度不要超过50 ℃，尽管针对特定的靶分子集合，要根据实验来确定最佳的清洗温度。

302

9. 在原先的试管中，用50 ml 2×SSC/0.1% SDS（预热到50 ℃）清洗杂交后的膜阵列，然后用2×SSC /0.1% SDS于50 ℃下旋转清洗两次，每次15 min。

   每次清洗步骤中用测量仪器监测阵列。是否需要更多的清洗要根据经验来确定。如果需要更加严格的清洗，将滤膜于65 ℃用1×SSC /0.1% SDS清洗5～15 min。

303

10. 将湿润的膜阵列在玻璃或者金属平板上排列整齐，用移液管在其上方滚动以除去所有多余的液体和小气泡。用塑料膜紧紧密封好滤膜，将密封膜在平板的背后粘牢。

    如果滤膜要重复使用（见下面的信息栏中），不要让它们干燥。为了使滤膜保持湿润，将一小片湿润的蘸墨纸放在板的底部。
11. 将微阵列曝光于磷像屏1～3天，然后在磷像仪（如Molecular Dynamics STORM磷像仪）中以50 μm的分辨率扫描成像屏。

    可以按照PART Ⅱ中介绍的方法对杂交结果进行分析。图4-2中给出了从一次典型的膜阵列实验中获得的结果。可能存在的问题和其相应的解决方法见图4-3中的介绍。注意可以从膜上除去标记的靶分子并重复利用，如下面的信息栏中介绍的那样。

附加方案：靶分子的清除和膜的重新利用

下面的一系列步骤介绍了从阵列上剥离去杂交的靶分子的方法，以便对膜进行重复利用。

附加材料

注意：标记有〈!〉的材料的正确处理方法参见附录3。

剥离溶液

0.4 mol/L NaOH〈!〉

0.1% SDS〈!〉

清洗溶液

0.2 mol/L Tris-HCl(pH 8.0)

1×SSC

0.1% SDS

方法

1. 为了除去标记的靶分子，将杂交后的膜在大约 200～300 ml 剥离溶液中于 65℃洗涤 2 次，每次 30 min 并剧烈摇动。
2. 将膜于室温下用清洗溶液洗涤 2 次，每次 10 min。
3. 让膜在空气中干燥，然后将它们对磷像屏曝光过夜，以确定剥离过程的效率。在信号发生较大程度的衰减之前，膜大约能够重复使用 3 次。

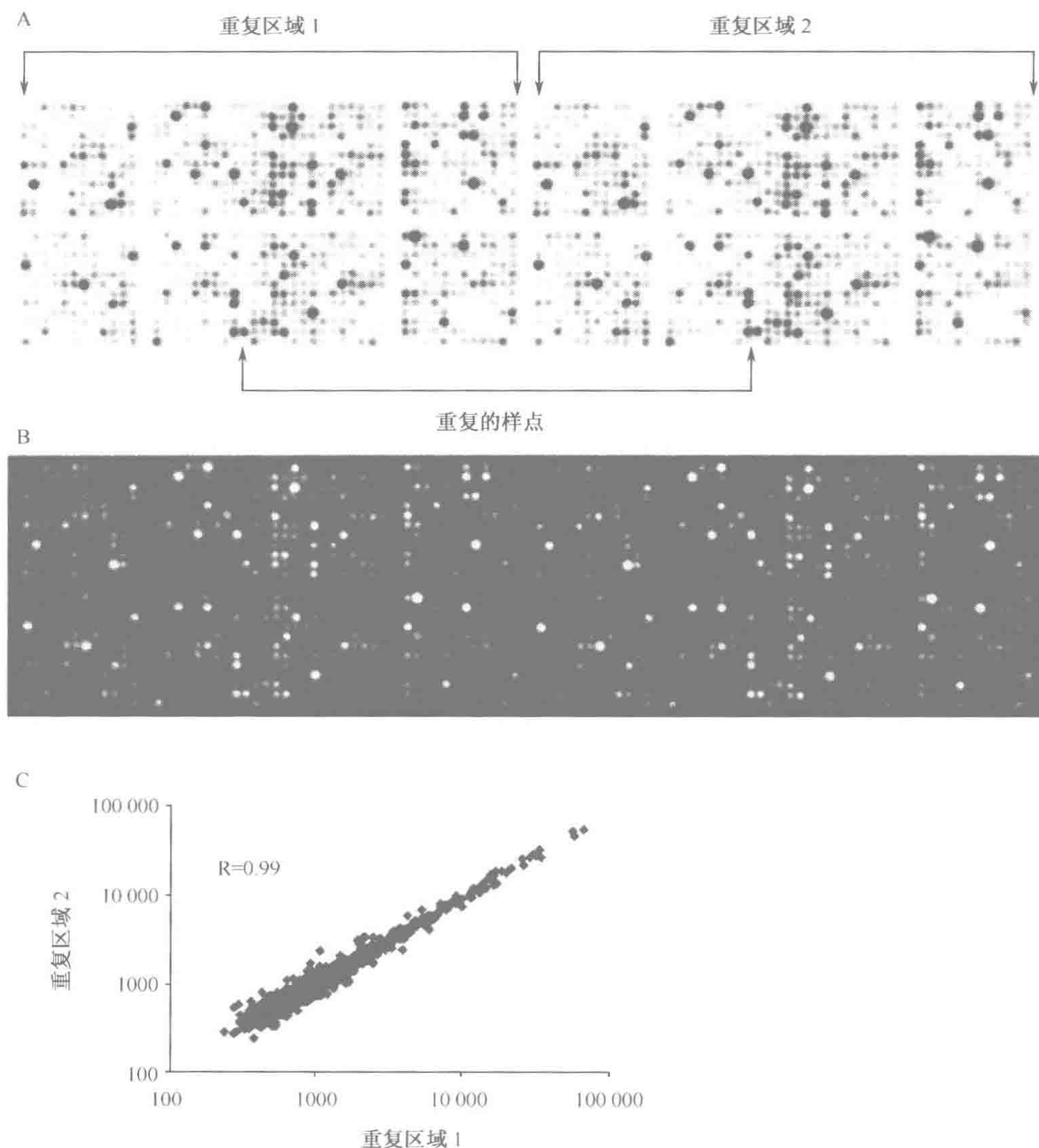


图 4-2　一个典型的包含了重复的膜阵列

A. 阵列与 5 μg 来源于 PC12 细胞系的总 RNA 杂交。该阵列宽大的重复区域中包含了 1152 种特定的 DNA 分子；B. 两幅放射性成像图像的伪彩色叠加；C. 总的非归一化强度的比较和重复区域之间的相关性（见彩图）

304

问题	解决方法
良好的图像	不需要
弱信号，低背景	延长曝光时间 重新标记RNA靶分子
弱信号，高背景	重新标记靶分子， 在新阵列上再次应用 标记的靶分子进行杂交
强信号，但带有 一致的杂交假象	重复实验 分析时剔除坏点， 回收靶分子
强信号，高的 一致性背景	重新清洗滤膜，回收 靶分子，对信号进行归 一化处理
点样模具相关 的点样假象	重新点样阵列，根据 每根点样针对信号进行归 一化处理
高背景，实 验彻底失败	回收RNA靶分子 重复整个操作过程， 实验中包括合适的对照 以识别产生失败的位置

图 4-3　膜阵列的疑难解答指南

305

PART Ⅱ　数据的分析和解释

许多软件程序可以应用于从原始的磷光图像中获得定量的数据，包括来自 Molecular Dynamics 的 ImageQuant（http://www.mdyn.com/products/ImageQuant），来自 Imaging Research 的 ArrayVision（http://www.imagingresearch.com），来自 Media Cybernetics 的 Array-Pro-Analyzer（http://www.mediacy.com/），来自 Raytest 的 AIDA Array(http://www.raytes.de/bio-img/bio_img.html) 和国立卫生院信息技术中心（http://cit.nih.gov）的 P-Scan。AtlasImage 是为分析 CLONTECH 膜（http://www.clontech.com）而专门设计的。采集原始数据的程序一般具有相似的功能并执行相似的任务。原始图像通常由许多较小的膜阵列组成。这些原始图像首先被分割成单个

阵列的图像，通过预先设置的方格图样或者自动的样点识别算法对子阵列的样点图样进行识别。强度的定量可以是无背景修正的，或者通过算法进行自动的背景修正。一旦获得强度的原始数据，就可以使用与分析玻片阵列相似的方法对他们进行分析。主要的差别在于测量的基本单位通常是原始强度的单一数值，而不同于双色阵列形式中实验组：对照组的比值。

与其他的阵列形式一样，归一化对于校正一系列样品之间由于标记或杂交差异而造成的变化起着关键的作用。没有一种归一化方法能解决所有的差异或者错误，并且不同的方法适合于不同的实验条件。因此，理解归一化的原因和进行恰当方法的选择是很重要的。

膜上非一致性的局部杂交异常有许多表现形式，并且一般在重复的实验中不会重复出现。因此，解决局部异常最有效的方法是重复实验。较小的局部异常可以进行标注，并从分析中略去。

背景的基线差异在重复实验中可以复制。一致性背景的影响可以简单地扣除掉，或者通过较复杂的归一化方法去除（Schuchhardt et al. 2000；Tanaka et al. 2000），这可应用于整体性归一化，逐根针进行背景校正，或者类似的局部背景校正。

一种作为替代的归一化方法是通过强度值的 Z 变换来实现（Vawter et al. 2001）。将强度的原始数据转换成对数值，从而减弱了偏差很大点的效应，并可以使得强度能够按照标准差（S. D.）为单位的方式给出（例如，Z 值的 1，2，3…表示了偏离阵列平均值 1，2，3…个标准偏差）。因为归一化函数特定于单个阵列而不适合于别的样品，所以 Z 值可用于样品之间和实验之间的比较。

一旦对原始数据进行了归一化或者变换处理，对来自膜阵列的数据进行进一步处理的绝大多数方法与那些应用于分析其他阵列形式的方法相类似（关于数据分析的详细讨论，见第 7 章，PART Ⅱ和 PART Ⅲ）。重复样点之间的方差和相关度量，相同 RNA 的重复标记（技术上的重复）以及重复性试验（生物学的重复）被用于监测阵列的质量和实验的重复性。对来自膜阵列数据进行的统计学评价、模式识别策略和可视化即使与其他阵列形式不完全一致，也是相似的（Eisen et al. 1998；Toronen et al. 1999；Tusher et al. 2001）。

参考文献

306

Andrews J., Bouffard G.G., Cheadle C., Lu J., Becker K.G., and Oliver B. 2000. Gene discovery using computational and microarray analysis of transcription in the *Drosophila melanogaster* testis. *Genome Res.* **10:** 2030–2043.

Barrett T., Cheadle C., Wood W.H., Teichberg D., Donovan D.M., Freed W.J., Becker K.G., and Vawter M.P. 2001. Assembly and use of a broadly applicable neural cDNA microarray. *Restor. Neurol. Neurosci.* **18:** 127–135.

Bertucci F., Bernard K., Loriod B., Chang Y.C., Granjeaud S., Birnbaum D., Nguyen C., Peck K., and Jordan B.R. 1999. Sensitivity issues in DNA array-based expression measurements and performance of nylon microarrays for small samples (erratum *Hum. Mol. Genet.* [1999] **8:** 2129). *Hum. Mol. Genet.* **8:** 1715–1722.

Eisen M.B., Spellman P.T., Brown P.O., and Botstein D. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci.* **95:** 14863–14868.

Friemert C., Erfle V., and Strauss G. 1989. Preparation of radiolabeled cDNA probes with high specific activity for rapid screening of gene expression. *Methods Mol. Cell. Biol.* **1:** 143–153.

Gress T.M., Hoheisel J.D., Lennon G.G., Zehetner G., and Lehrach H. 1992. Hybridization fingerprinting of high-density cDNA-library arrays with cDNA pools derived from whole tissues. *Mamm. Genome* **3:** 609–619.

Lennon G.G. and Lehrach H. 1991. Hybridization analyses of arrayed cDNA libraries. *Trends Genet.* **7:** 314–317.

Lockhart D.J., Dong H., Byrne M.C., Follettie M.T., Gallo M.V., Chee M.S., Mittmann M., Wang C., Kobayashi M., Horton H., and Brown E.L. 1996. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Biotechnol.* **14:** 1675–1680.

Maier E., Meier-Ewert S., Ahmadi A.R., Curtis J., and Lehrach H. 1994. Application of robotic technology to automated sequence fingerprint analysis by oligonucleotide hybridization. *J. Biotechnol.* **35:** 191–203.

Sargent T.D. and Dawid I.B. 1983. Differential gene expression in the gastrula of *Xenopus laevis. Science* **222:** 135–139.

Schena M., Shalon D., Davis R.W., and Brown P.O. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270:** 467–470.

Schuchhardt J., Beule D., Malik A., Wolski E., Eickhoff H., Lehrach H., and Herzel H. 2000. Normalization strategies for cDNA microarrays. *Nucleic Acids Res.* **28:** E47.

Takahashi N. and Ko M.S. 1994. Toward a whole cDNA catalog: Construction of an equalized cDNA library from mouse embryos. *Genomics* **23:** 202–210.

Tanaka T.S., Jaradat S.A., Lim M.K., Kargul G.J., Wang X., Grahovac M.J., Pantano S., Sano Y., Piao Y., Nagaraja R., Doi H., Wood III W.H., Becker K.G., and Ko M.S. 2000. Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray. *Proc. Natl. Acad. Sci.* **97:** 9127–9132.

Taniguchi M., Miura K., Iwao H., and Yamanaka S. 2001. Quantitative assessment of DNA microarrays—Comparison with Northern blot analyses. *Genomics* **71:** 34–39.

Toronen P., Kolehmainen M., Wong G., and Castren E. 1999. Analysis of gene expression data using self-organizing maps. *FEBS Lett.* **451:** 142–146.

Tusher V.G., Tibshirani R., and Chu G. 2001. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci.* **98:** 5116–5121.

Vawter M.P., Barrett T., Cheadle C., Sokolov B.P., Wood III W.H., Donovan D.M., Webster M., Freed W.J., and Becker K.G. 2001. Application of cDNA microarrays to examine gene expression differences in schizophrenia. *Brain Res. Bull.* **55:** 641–650.

Whitney L.W., Becker K.G., Tresser N.J., Caballero-Ramos C.I., Munson P.J., Prabhu V.V., Trent J.M., McFarland H.F., and Biddison W.E. 1999. Analysis of gene expression in mutiple sclerosis lesions using cDNA microarrays. *Ann. Neurol.* **46:** 425–428.

因特网资源

http://cit.nih.gov Center for Information Technology homepage, National Institutes of Health
http://www.clontech.com BD Biosciences Clontech homepage
http://www.resgen.com ResGen, trademark of Intvitrogen Corporation
http://www.imagingresearch.com Imaging Research Inc. homepage
http://www.mdyn.com/products/ImageQuant Molecular Dynamics image analysis software
http://www.mediacy.com MediaCybernetics homepage
http://www.raytest.de/bio-imag/bio_imag.html Raytest's bioimaging products and applications

第5章 组织显微切割

Carolyn J. M. Best，Isabel M. Leiva，John W. Gllespie，Rodrigo F. Chuaqui，Michael R. Emmert-Buck
（病理遗传学研究室，病理学和泌尿肿瘤学实验室，国家癌症研究所，贝塞斯达，马里兰州，20892）
Renate Burgemeister and Karin Schütze
（P. A. L. M 微激光技术公司，贝恩里德市，德国）

导言

308

组织是由多种相互作用的细胞种群按照复杂的三维空间排布而成的，每种细胞表型反映了一种不同的信使 RNA（mRNA）和蛋白质的表达谱。生物学研究的一个重要目的就是研究样品中特定类型细胞的分子表达谱（图 5-1）。然而，研究中感兴趣的细胞

可能在样品中只占到总细胞的一小部分，并且相互之间可能会被组织的其他成分分隔开来。重新获取这些在空间区域上难以接近的细胞，涉及使用能够将靶细胞与相邻细胞分离开的显微切割技术。在这些技术被开发出来之前，免疫组织化学和原位杂交技术是用于测量组织中特定细胞内蛋白质和 mRNA 水平的主要工具。尽管这些方法是非常有用的，但是局限于分析单个的基因或者是基因产物，不适用于对基因组重排或者对 mRNA表达的整体性研究。

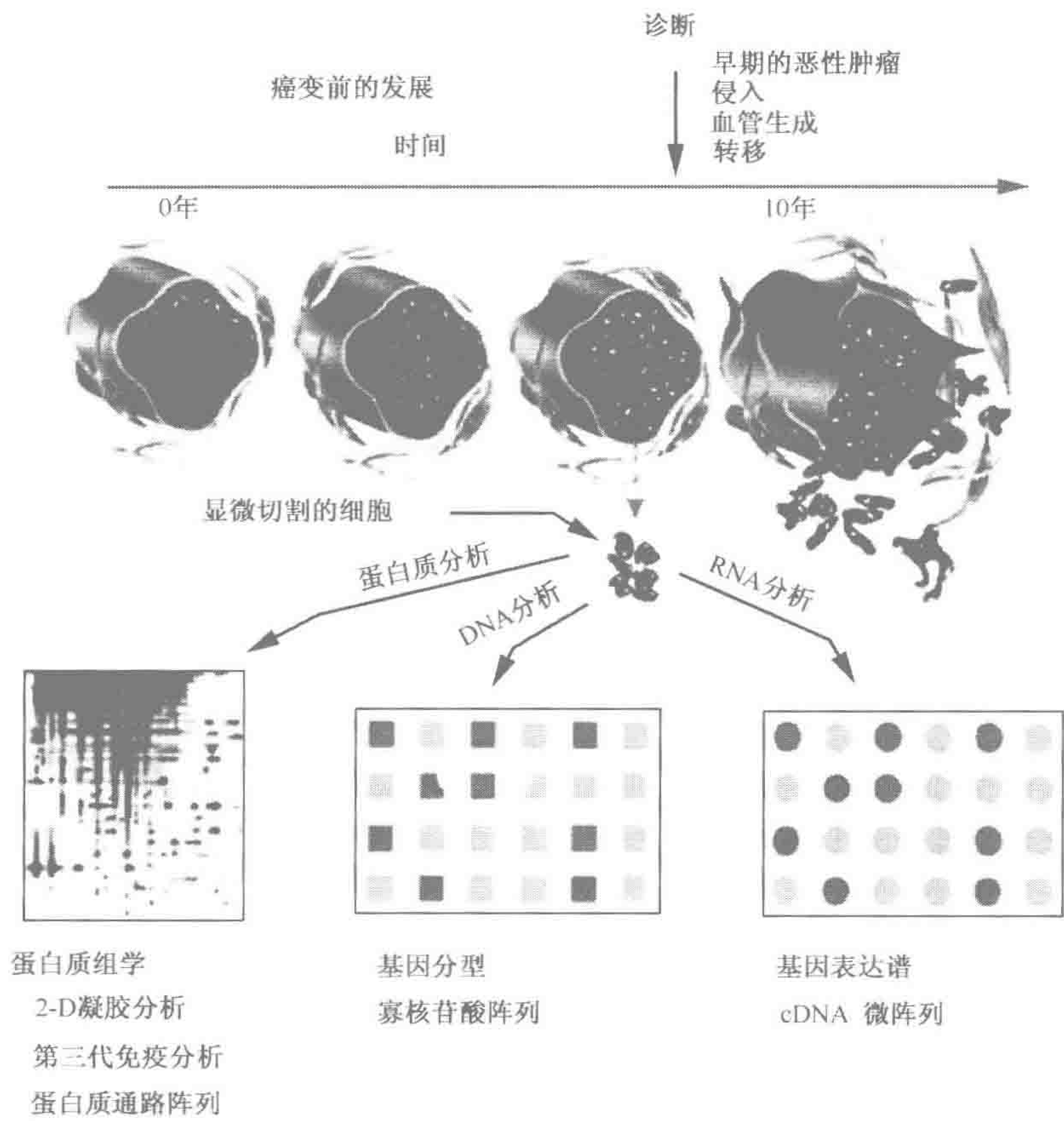


图 5-1　应用组织显微切割研究与肿瘤发展相关的分子事件

可以对单个患者组织中的蛋白质分子谱（2-D 凝胶分析或者免疫分析）、DNA 基因分型（如突变分析和杂合性缺失）或者 RNA 的分子谱（如 cDNA 文库和微阵列）进行比较。正常的上皮细胞、基质、癌变前损害、入侵性肿瘤细胞和炎症性细胞都可以进行取样和作比较

309　显微切割方法

直接在显微观察下从非均匀的组织样品中分离特定的细胞种群在理论上是简单的，但在实际应用中，这种方法具有技术上的挑战性，并且经过了发展阶段的演变（表 5-1）。

表 5-1　组织切片显微切割方法概要

	徒手组织显微切割[a]	MOMeNT[a]	LCM[a]	LPC[a]	LMPC[a]
原理	使用或者不用倒置显微镜，通过灭菌的解剖针或刀片来获取组织切片上大块区域的徒手切割	采用聚乙烯薄膜作为支持膜，从组织切片上切取成群细胞的徒手组织切割	通过红外激光产生的热量，选定的单个或成群细胞从组织切片转移到热塑性薄膜上	非接触式弹射过程：由 UV-A 激光产生的气体压力将单个或成群的细胞从目标一侧提升；需要的细胞通过“非接触式”的激光弹射来回收；不产生热量	非接触式的显微切割过程；UV-A 激光勾勒将选定的标本与周围区域分离开来，并在该组织区域内选择性地除去不需要的细胞
回收样品的方法	手工俘获（灭菌的解剖针或刀片）或者显微操作器；转移到 PCR 管内	借助于膜使得手工俘获（灭菌的解剖针、刀片、镊子或者显微操作器）变得容易	将附着在热塑性转移膜上的细胞直接收获到 LCM 盖子上	激光直接弹射到收集表面上（充满缓冲液的盖子、排盖或者微量滴定板）	激光直接弹射到收集表面上（充满缓冲液的盖子、排盖或者微量滴定板）
回收样品的最小尺寸	≥ 50～100 μm	50～100 μm	>7.5 μm	<1 μm	<1 μm
最佳应用	回收大范围的、均质的细胞	回收大范围的、均质的细胞	一群或者是少数几群细胞（5～50）	单个的或者是一小群的细胞（50～100）	染色体，核，单个细胞，成群细胞（1～10 000），活细胞
标本的制备	FFPE，FF	FFPE，FF	FFPE，FF，细胞涂片；细胞抹片	FFPE，FF，细胞涂片；细胞抹片	FFPE，FF，细胞涂片，细胞抹片，细胞培养物
制备时间	5～10 min，> 10^4个细胞	5～10 min，1～10 个细胞	< 10 s，最多 30 个细胞	< 1 s，1～30 个细胞；< 10 min，最多 10 000 个细胞	< 1 s，1～30 个细胞；< 5 min，最多 10 000 个细胞
仪器费用	低	低	高	高（自动）	高（自动）
优点	简便快捷	分离特定的细胞种群的快速方法，且污染最少	保留 RNA 完整性的一种非常快捷的方法；直接观察确证俘获的细胞	非常快捷且便于操作的方法，保存了 DNA、RNA 和蛋白质的完整性	非常快捷且便于操作；高精度地回收细胞种群；能够从培养物中回收活细胞；直接观察确证
缺点	高污染风险；不适合于病理改变的标本或者多种类型相互混合细胞构成的组织	组织切片制备复杂，仅适用于膜支持的组织切片；不能用于常规处理的玻璃载玻片	需要转移膜；不能用于活细胞的俘获	弹射过程中使得形态学变得不完整	需要使用膜支持的标本以保持形态学特性

根据 Walch 等（2001）改编而成。

a 缩写：（FFPE）经甲醛固定、石蜡包埋的组织；（FF）新鲜冰冻组织；（MOMeNT）膜支持天然组织的显微切割；（LCM） 激光俘获显微切割；（LPC） 激光压力弹射；（LMPC） 激光显微切割和压力弹射。

- **徒手显微切割**。早期的研究需要用解剖刀片挖取切片以便从冰冻的组织学载玻片上粗略地去除组织，并得到感兴趣的细胞。此后，徒手或者在显微操作器导引下的带有黏性针头解剖针应用改善了显微切割的准确性和可靠性（Going and Lamb 1996）。
- **放置在膜上的自然组织的显微切割（MOMeNT）**。MOMeNT 是对徒手显微切割的改进，涉及将组织切片放置于黏附在载玻片上的 6 μm 厚的聚乙烯膜上。薄膜用作使全部组织集中在一起的支持物，这样可以使用解剖刀徒手环切大块的组织，并用镊子进行获取（Böhm and Wieland 1997）。后来这种方法经过改进，改用较薄的金属箔片，这样可以使用激光切割勾勒出一系列的靶细胞群，然后使用解剖针或者镊子进行手工获取（Böhm et al. 1997；Gjerdrum et al. 2001；Ling et al. 2001；Specht et al. 2001；Walch et al. 2001；Fink et al. 2002）。
- **使用 UV 照射消蚀细胞**。一种被称为选择性紫外照射分离（SURF）技术的进步使得不想要的细胞可以被 UV 照射冷消蚀。在这个方法中，UV 照明器（在 UV-C 范围，波长 302 nm）照射整张载玻片，引起单链缺口和对细胞活力的破坏。感兴趣的细胞使用墨水进行防护而免受破坏，随后使用显微操作器取出这些细胞（Shibata et al. 1992）。
- **激光俘获显微切割（LCM）**。LCM 是一种基于激光的显微切割方法，它使用了一台整合到标准的倒置式显微镜照明光路中的红外激光器。来自红外激光器的光脉冲通过一个透明的硬塑料盖照射，在塑料盖的表面衬有一层热塑性的乙烯-乙酸乙酯薄膜。这种专门设计的注塑盖子，在沿着周长的方向有支撑的轨道，所以当它放置在按照常规方法制备的组织切片表面上时，在热塑性薄膜和组织之间产生一个 12 μm 的狭缝。在显微镜下观察组织切片，当靶细胞位于这层转移膜的下方时，激光二极管被触发。激光脉冲的加热效应（980 nm）使得相当于激光焦斑尺寸大小（60 μm、30 μm、15 μm、7.5 μm）的膜区域发生局部的熔化。熔化的薄膜流到组织切片上迅速冷却，与靶细胞紧密地结合。当膜移开时，研究者所选择的组织区域从切片中获取。空间分布复杂的靶细胞通过在不同位置上重复这一过程来获取，直到足以用于分析的大量细胞结合到薄膜上为止。
- **激光压力弹射（LPC）**。“非接触式的”LPC 利用聚焦的 UV-A 激光直接转移单个细胞或者成群的细胞——没有机械的接触——到合适的收集表面（如装有缓冲液的标准离心管的盖子）（Fink et al. 1998；Schütze and Lahr 1998；Lahr 2000；Lehman et al. 2000；Glöckner et al. 2001；Imamichi et al. 2001；Scheidl et al. 2002；Westphal et al. 2002）。俘获的样品通过简单的离心可以很容易地转移到试管中。脉冲的低能量 337 nm 氮激光器通过标准显微镜的落射荧光通道整合到系统中，并通过物镜被聚焦。弹射的物理作用力被认为是来自于因受到激光焦面上巨大的光子流引发的消蚀而在靶区域处产生的高压的气体。根据样品本身的特性，显微切割可以在弹射之前完成，这一过程被称为激光显微切割和压力弹射（LMPC）。这个完全自动化的方法使得快速制备纯粹和均一的样品变得容易。激光束在导引下沿着所希望的细胞区域的外围切割，这样使他们和相

邻的组织成分分开。也可以将激光束引入到组织的特定区域以消除在精确选定的区域中不需要的那些材料。聚焦的激光束引发局部的光分解，并在其狭小的焦点处消蚀不需要的组织。

310

目前所使用的主要显微切割方法的原理和特性归纳于表 5-1 中。有关激光俘获显微切割（LMC）和 LMPC 的更多详细信息，分别参见方案 2 和方案 10 以及本章中 PART Ⅰ和 PART Ⅱ的导言部分。

组织处理的重要性

用于显微切割的材料有许多种生物学来源，它们常常是患者手术切除的、被送到组织病理学实验室中的活检样品（图 5-2）。乍看上去，这些样品如何获得和处理似乎与那些分析分子表达谱数据集的生物学家和生物信息学家无关。然而，当样品从患者身体切除后，在处理之前或者处理过程中，细胞的分子表达谱可能会发生改变。影响 RNA 的回收和分析的因素有：样品手术切除和在病理实验室中处理之间的时间间隔、固定试剂和包埋介质的类型、固定时间的长度以及处理温度。所有这些变量都会对使用 LCM 和 LMPC 从组织得到的样品中抽提出的大分子的产量和质量产生显著的影响（如 Goldsworthy et al. 2000；Fend et al. 1999a，b；Kleeberger et al. 2001；Scheidl et al. 2002）。在计划进行微阵列实验时，样品从外科医生送给病理学家，最后到实验室中，因而建立一个系统以优化样品的采集和处理是非常重要的。下面几点要加以考虑：

312

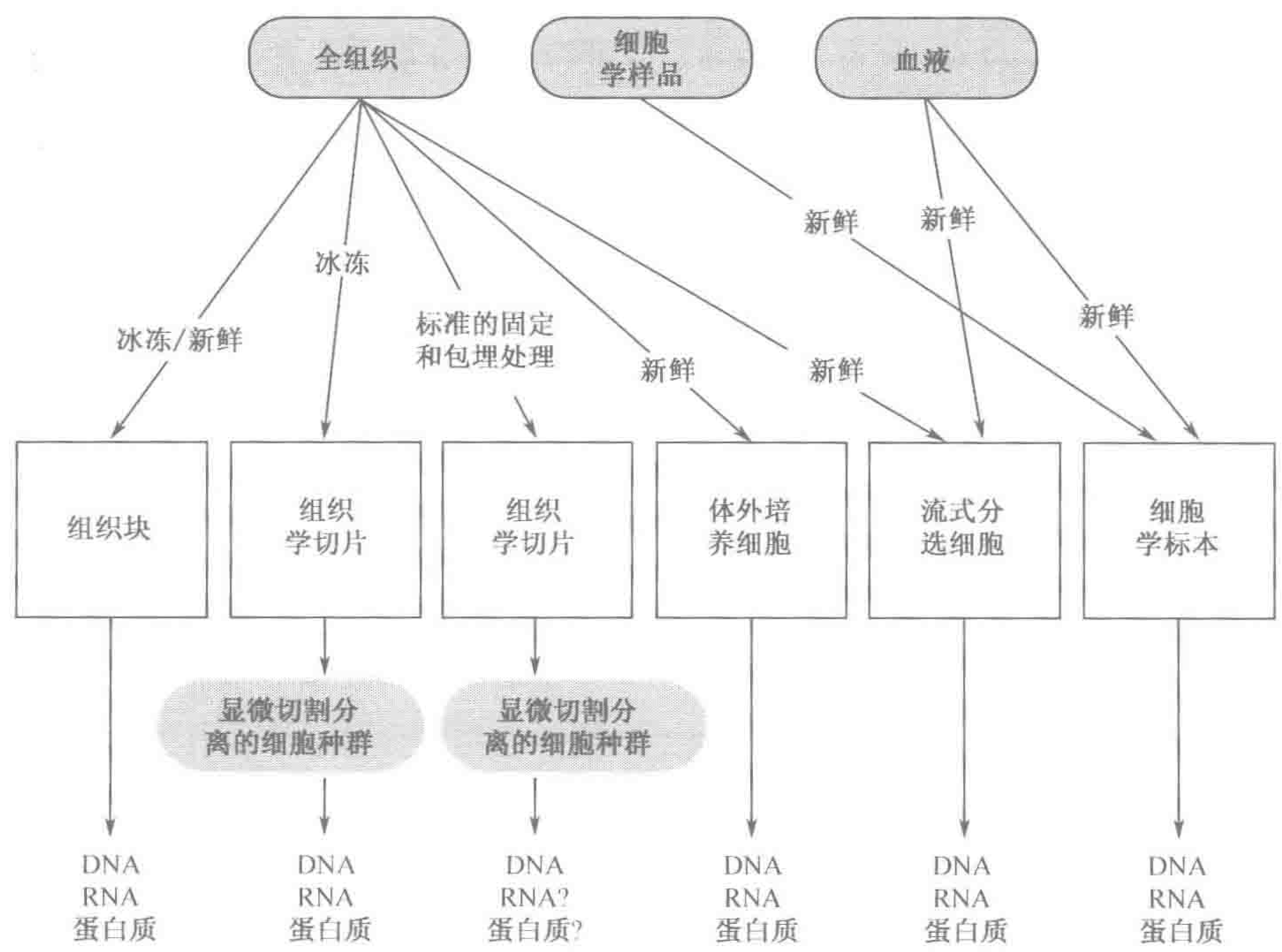


图 5-2 样品处理对分子分析结果的影响

313

对于每种人体样品，考虑处理方法时必须根据所选定的后续分析作决定。相对于后续的分子分析而言，每种处理方法都有特定的优势和不足，研究人员必须选择与其研究的特定目标相适应的合适方法。关于每种处理方法在使用上的建议，参考 http：//cgap-mf. nih. gov/BtoB/BtoBMolAnal. html.

• 可能获得的组织样品的种类和数量。
• 如何得到患者的许可。
• 如何以及何时获取样品。
• 病理学实验室中将如何处理样品。
• 在整个处理过程中，每一步如何跟踪、记录样品，有必要的话，还要隐匿样品的身份信息。
• 在样品到达微阵列实验室中保管时，如何保存和处理样品。

一个非常重要的问题是病理学实验室中处理组织时使用的方法。在日常应用中，病理学家有三个直接的选择：冰冻活检样品、保持样品新鲜或者进行固定和包埋。许多外科手术和活检样品在10%中性缓冲的甲醛溶液中固定，然后在石蜡中包埋。包埋过程从使用一系列浓度不断增加的乙醇进行脱水开始，随后在二甲苯中“透明”处理，使石蜡可以完全地渗透到组织块中。对石蜡块切片并转移到载玻片上，用于染色、显微观察以及最后制备显微切割的细胞种群。

病理学样品通常用甲醛固定，这是因为标本的结构比冰冻处理保存得更加完好。在诊断完成后，很多病理学实验室要保留一些包埋的组织块作为档案储存，这些存档材料在需要时可以用于追溯，如确诊或者对诊断进行评估。但是，在告知患者并得到同意，且获得 Institutional Human Research Ethics Committee 的许可之后，一部分不需要用于诊断或存储的活检材料可以提供给研究项目。通常，来自甲醛固定和石蜡包埋的组织块，用薄片切片机切成厚度为 5 μm 的切片裱贴在标准载玻片上，用苏木精、曙红、甲基绿或快红进行染色（Burton et al. 1998）后进行显微分析。在显微切割之前，组织切片有时进行免疫组化染色以鉴别特定类型的靶细胞（Fend et al. 1999a，b；Noack et al. 1999；Buerger et al. 2000a，b；Kaserer et al. 2000；Persson et al. 2000；Gjerdrum et al. 2001；Ling et al. 2001）。虽然在甲醛中固定的组织在形态学保持上要比冰冻材料好，但将经过甲醛固定，石蜡包埋的组织标本用作高通量表达阵列的样品时有很多问题，主要是由于在固定过程中蛋白质和 RNA 发生了交联。

冰冻切片目前被认为是显微切割最佳的材料，因为从它们可以得到高质量的 DNA、RNA 和蛋白质（Goldsworthy et al. 1999）。将新鲜的、经过再次切割的组织在一种包埋复合物，如 OCT 中冰冻，然后在冰冻切片机上得到冰冻切片（Hiller et al. 1996；Ponten et al. 1997；Kaserer et al. 2000；Persson et al. 2000；Stoehr et al. 2000；Ling et al. 2001）。不幸的是，在这个过程中，细胞组织学有时被破坏，使得靶细胞群的识别变得困难。显然，为了满足病理学家和分子生物学家的需求，必须要研究改进的样品固定和包埋策略，提供一个能被病理学家所接受的组织学结构分辨水平，有效地显微切割选择的细胞，并保持组织内转录物的完整性。研究人员应当花费一些时间来建立下面的系统：

• 能够回收足够量的 RNA 用于实时 PCR 分析、体外扩增，生成用于表达谱分析的荧光或者放射性标记的靶分子（Luo et al. 1999；Glöckner et al. 2000；Lehmann et al. 2000；Nagasawa et al. 2000；Ohyama et al. 2000；Bertheau et al. 2001；Hoffmann et al. 2001；Specht et al. 2001；Scheidl et al. 2002）和cDNA文库的构建（Peterson et al. 1998）

- 获得标本的图像，能清楚地识别出细胞类型。图像质量越佳，样品的细胞病理学与下游的分子分析之间的联系越紧密（Emmert-Buck et al. 1996，2000；Bonner et al. 1997；Simone et al. 1998；Banks et al. 1999；Dietmaier et al. 2000；Vona et al. 2000；Gjerdrum et al. 2001；Westphal et al. 2002）。显微切割前后组织切片的数字图像为实验提供了不可或缺的记录。

开始实验前最好阅读 Goldsworthy 等（1999）的论文。这些作者比较了使用 LCM 从小鼠肝脏切片获得的样品中，各种固定方法对细胞形态学和 RNA 扩增的影响。如同预期的那样，从石蜡包埋材料中得到的 RNA 质量差。乙醇（70%）是理想的形态学和显微切割首选固定剂。用冰冻组织抽提效果最佳。在人的前列腺样品中也得到类似的结果（Gillespie et al. 2002）。尽管如此，实验人员应当尝试多种固定方法，因为一种方法不可能对所有的组织和条件来说都是理想的。

致谢

我们非常感谢 Arcturus，Inc. 和 P. A. L. M. Microlaser Technologies 的科学家们，感谢他们在本章写作和审阅时所提供的帮助：James Stanchfield，Steve Kunitake，Anna Mennis 和 Rachel Goodrich。

PART Ⅰ 徒手显微切割和激光俘获显微切割 314

Carolyn J. M. Best，Isabel M. Leiva，John W. Gillespie，Rodrigo F. Chuaqui，Michael R. Emmert-Buck
（病理遗传学研究室，病理学和泌尿肿瘤学实验室，国家癌症研究所，贝塞斯达，马里兰州，20892）

导言

在过去的十年中，出现了分子谱分析这一充满活力的新学科，能够在整体上观察研究各种类型细胞和疾病过程中的 mRNA、蛋白质的表达谱，以及 DNA 的分子变化。分子谱分析将来自不断增长的人类基因组计划的遗传数据库的信息和新近发展起来的表达分析技术集成在一起。这种新方法使得对大量仍未知晓的人类基因型和表型之间联系的理解变得容易，并有望发展用于临床干预的诊断和治疗的新靶标。

测序和扩增方法的最新进展使得对来自少量的、离散的生物样品（如细胞学涂片和组织活检）的 DNA 和 RNA 的分析变得容易。然而，这些方法的成功与否和精确性高低取决于所研究细胞种群的完整性和纯度。手工和机械方法分离感兴趣的离散细胞种群，尽管在一定范围内的项目中是有效的，但是对于广泛的研究项目而言，往往非常耗费时间和人力。

新型的显微切割技术，依赖于使用集成到标准的显微镜中的低能量红外激光，能够从标准的组织学切片中回收特定的细胞种群。LCM 最早在 NIH（national institutes of health）的 NICHD（national institute of child health and human development）和 NCI（national cancer institute）中开发出来（Emmert-Buck et al. 1996，1997；Bonner et al. 1997）。Arcturus 公司是一个致力于设计和开发分子诊断技术的公司，作为合作研究和开发协议的一部分而与 NIH 一起合作，将 LCM 发展为商品化的系统。有关 LCM 技

术、应用和方案的详细信息，见 Arcturus 的网站 www. arctur. com。

NCI 的病理遗传学所致力于基于组织显微切割方法的分子谱分析（例如，见图 5-3）。LCM 被用于癌症基因组解剖项目（cancer genome anatomy program，CGAP），从正常到疾病状态对细胞的发展进行分类。这项技术适用于任何的疾病过程，通过组织取样，如癌变前的肿瘤病变、多发性硬化、动脉硬化和 Alzheimer 病而实现。更多的信息，见分子表达谱项目网站 http://cgap-mf. nih. gov/index. html。

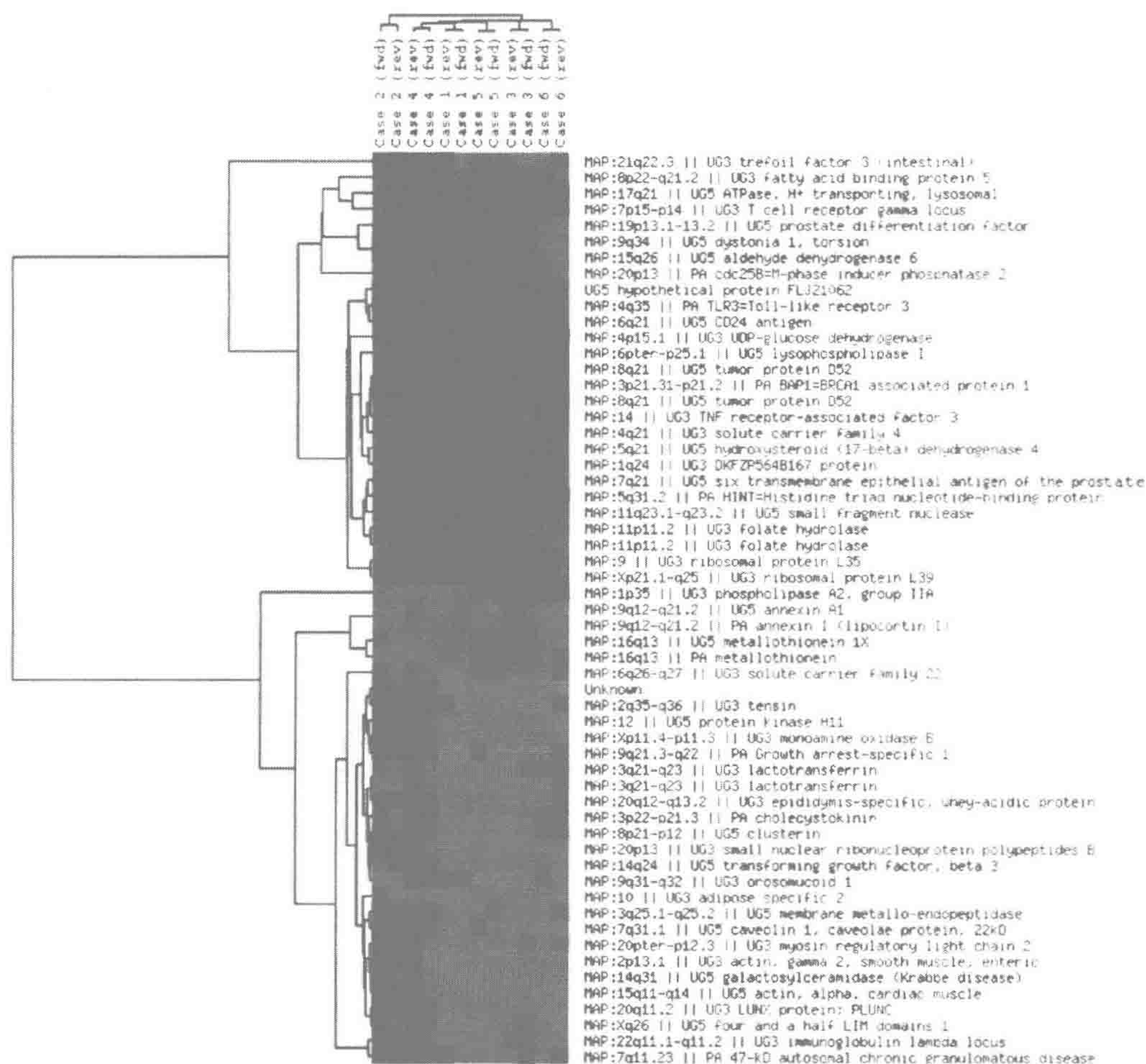


图 5-3　前列腺癌的微阵列分析

来自 6 个高度癌变的前列腺肿瘤和一个从正常的前列腺组织合并得到的对照样品的总 RNA 被用于生成荧光标记的第一链 cDNA 靶分子。探针重复杂交（正向和反向荧光反应）到 6.4K 人类 cDNA 微阵列（NCI Array Facility，Gaithersburg，Maryland）。肿瘤样品紧密地集中在一起，并且有 52 个基因被鉴别出在大多数的肿瘤中显示出一致性的变化（见彩图）

PART Ⅰ的一系列方案提供了当前应用的各种组织显微切割方法的详细情况和应用。

- 方案 1 中介绍了使用 30 号针和倒置显微镜对染色的组织样品进行徒手显微切割分离。

- 方案 2 给出了使用 LCM 对组织样品进行分离。
- 方案 3 介绍了对特定的、选择的细胞种群进行免疫染色，为 LCM 作准备。
- 方案 4 给出了从显微切割的样品中回收 RNA 的各种方法。

315

有关 LCM 和激光切割显微镜更多的方案、信息和参考资料（请见 Arcturus 网站：www.arctur.com），第 3 章中（方案 13～方案 18）的一系列方案介绍了对显微切割样品更进一步的下游处理，以便用于微阵列试验。第 3 章的方案 20 介绍了如何建立和处理与微阵列的杂交。

方案 1

316

徒手显微切割

徒手显微切割和后续的分子分析可以在含有标准的、经过甲醛固定的、5 μm 厚的组织切片中，并在苏木精、曙红或极为少见的甲基绿、快红染色的载玻片上进行。使用倒置式显微镜，组织切片用 30 号针进行分离。虽然很多研究人员选择使用安装了长工作距离物镜（10×和 50×干物镜）的显微镜，我们未发现必须使用安装了特殊物镜的显微镜。使用普通物镜，在实际应用中可使用的最高放大倍数为 10×，它能够实现相对精确的徒手显微切割。使用观察组织需要的最少量的染色进行显微切割可以显著改善大分子的回收。由于是从没有盖玻片的组织切片进行显微切割，样品和空气的折射率相差很大。这使得光轴上图像发生扭曲并且光散射较为严重，常常产生“暗”的图像。将染色浓度减少到其标准浓度的大约 10%有助于改善图像质量和分子的回收。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

蛋白酶抑制剂混合片剂 Complete miniprotease inhibitor cocktail tablets(1 836 153，Roche)

对于所有的蛋白质分析，除了二甲苯以外，每 10 ml 各种试剂要溶解一片蛋白酶抑制剂混合片剂。

去离子水

曙红 Y 溶液（Sigma）

乙醇（70%，95%，100%）

据报道，有些批次的乙醇（100%或 200 标准强度）中含有会使微阵列试验中产生很深荧光背景的污染物质，95%的乙醇不存在这个问题。为避免这个问题，可以用 95%的乙醇或已鉴定证明不会产生高荧光背景的无水乙醇。Warner-Graham 公司（Cockeysville,

Maryland）出售的无水乙醇（USP 级乙醇，64-17-5）效果很好。

显微切割样品的抽提缓冲液（选择适合于显微切割样品的缓冲液，也就是说 DNA、RNA 或蛋白质分析：见步骤 15）

DNA 抽提缓冲液

50 mmol/L Tris-Cl(pH 7.6)

1 mmol/L EDTA(pH 8.0)

1% Tween-20

2 mg/ml 蛋白酶 K

317

RNA 抽提缓冲液

RNA 可以使用很多商品化试剂盒中的任何一种从显微切割的细胞中回收。通常，细胞在含有β-巯基乙醇的、标准的异硫氰酸胍的缓冲液中匀浆，然后进行 RNA 的纯化，如 QIAGEN的 RNeasy 试剂盒。大多数情况下，当 RNA 仍然结合在硅胶膜上时，在清洗步骤之前使用 DNase Ⅰ消化以除去残留的基因组 DNA。更详情的信息，见第 3 章，以及本章的方案 1 或方案 4。

蛋白质抽提缓冲液

如果将蛋白质用于电泳分析，将细胞在下面的缓冲液中裂解：

8 mol/L 尿素

2 mol/L 硫尿

2%(*m/V*) CHAPS

1%(*m/V*) 二硫苏糖醇（DTT）〈!〉

0.8% Pharmalyte 两性电解质载体（pH 3～10）

更多信息，见 Banks 等（1999），Emmert-Buck 等（2000）和 Ornstein 等（2000）。

如果从冰冻切片中回收和分析天然蛋白质，见本方案结束部分**替代方案：冰冻组织样品的显微切割**中的导言部分。

甘油（超纯级，GIBCO）

苏木精溶液（Mayer，Sigma）

二甲苯（混合）（ACS 试剂，Sigma）

细胞和组织样品

组织切片

从经甲醛固定、石蜡包埋的组织切下的切片裱贴在载玻片上，在室温或者低于室温下保存。直到显微切割之前，不要给切片脱蜡。包埋在低熔点聚酯中的切片在 4℃下保存。冰冻切片在－80℃或者更低的温度下保存。

专用设备

倒置显微镜（标准的）

显微切割工具（装在注射器上的 30 号针）

方法

载玻片的处理

对于石蜡包埋的切片，从步骤 1 开始；对于冰冻包埋切片，从步骤 4 开始；对于聚

酯包埋的切片，将载玻片在无水乙醇中浸泡两次，每次 5 min（使切片脱蜡），然后从步骤 3 开始。

1. 为使组织切片脱蜡，将载玻片在新鲜的二甲苯中浸泡 5 min。
2. 将载玻片转移到新鲜的二甲苯中，浸泡 5 min。
3. 将载玻片转移到无水乙醇中，浸泡 30 s。
4. 将载玻片转移到 95%的乙醇中，浸泡 30 s。
5. 将载玻片转移到 70%的乙醇中，浸泡 30 s。
6. 将载玻片转移到去离子水中，浸泡 30 s。
7. 将载玻片转移到 Mayer 苏木精溶液中，浸泡 30 s。

318

8. 将载玻片转移到去离子水中 15 s；重复一次。
9. 将载玻片转移到 70%的乙醇中，浸泡 30 s。
10. 将载玻片转移到曙红 Y 中，浸泡 15 s。
11. 将载玻片转移到离子水中，浸泡 30 s；重复一次。
12. 将载玻片转移到含 3%甘油的去离子水中，浸泡 30 s。

    在 3%甘油中浸泡，特别有助于准备用于显微切割的组织，因为它使得组织的脆性下降，切割的组织碎片容易收获。

13. 将玻片从 3%甘油溶液中取出，轻轻摇动以除去甘油/水层。

    从甘油/水中取出玻片时，确保除去覆盖在载玻片上的一薄层液体非常重要。在这层液体存在的情况下进行切割会导致组织碎片的扩散，有可能“污染”样品。此外，在液体下方进行切割会产生大块的组织剥落物，不容易在抽提缓冲液中匀浆。

切割

如果可能的话，在接下来的 5～10 min 内进行显微切割。在这段时间内，组织是干的，但仍旧保持整体上的柔软。如果切割花费一些时间，组织会逐渐变脆，而且当针接近组织时，切割下来的碎片可能会被排斥开。如果组织变得过度干燥，将其在 3%甘油/水溶液中重新浸泡 1～2 min。

14. 通过显微镜观察组织时，使用一根安放在注射器上的 30 号针轻轻刮取感兴趣的细胞种群。切割的细胞会与载玻片分离，并形成一小团暗色的组织，通过静电吸引收集到针上。可以同时收获一些小的组织碎片。将第一块组织碎片收集到针头上，以帮助收获后续的组织。

    为使进行切割操作的手保持稳定，将肘部支撑在与显微镜载物台具有相同的高度并靠近载物台附近的固体表面上。把进行切割的手的尺骨靠在显微镜载物台上，有助于移动解剖针进入显微镜的视野，并位于组织上方数毫米。用这种方法，进行切割的手臂和手都可以靠在固体支撑物表面。

15. 小心地将带有收获组织碎片的针尖放入含有 100 μl 适合于所研究组织样品的抽提缓冲液的 1.5 ml 微量离心管中。轻轻晃动试管以保证组织从针尖上脱离。

    按下注射器的柄向抽提溶液中注入空气泡，这有助于组织从针上脱离，并防止碎片残留在针管里。

16. 重复步骤 14 和步骤 15，继续收获细胞或组织样品。

要收集的细胞或组织样品的数量取决于组织类型以及实验的目的。在前列腺组织中，我们建议收集5000个细胞用于RNA扩增（如在微阵列分析之前）和100 000个细胞用于双向蛋白质凝胶电泳。

319

替代方案：冰冻组织样品的显微切割

如果要从冰冻切片以天然形式回收蛋白质，将冰冻组织切片直接放在琼脂糖包被的载玻片上有助于保持稳定性。琼脂糖凝胶可以在合适的缓冲液中制备或浸泡于其中，该缓冲液在显微切割前或者切割过程中将浸透冰冻切片。缓冲液的pH、盐浓度、蛋白酶抑制剂等应适合于所需要靶蛋白质的性质。琼脂糖包被的载玻片也可以用于回收mRNA的显微切割。

附加材料

在合适缓冲液中的琼脂糖溶液（2%），加热到45℃

冰冻切片机

载玻片（标准的未包被载玻片）

方法

1. 要制备用于显微切割的载玻片，向标准的未包被玻璃载玻片上加入200 μl温热的2%琼脂糖溶液，盖上玻璃盖玻片，让凝胶固化。
2. 从载玻片上移去玻璃盖玻片，立即将新鲜切下的组织切片从冰冻切片机转移到琼脂糖包被的载玻片上。
3. 使用30号针将组织"梳理"开来。

   组织将仍旧浸透在来自凝胶的液体里，并能够被轻轻分开，沿着组织层面分离（例如，基质和上皮细胞相互之间能够很容易地分开。）
4. 轻轻地从载玻片上回收切割的组织，或者，作为替代方法，用针切开琼脂糖，并将琼脂糖和组织碎片同时收获。

320

方案2

激光俘获显微切割

在激光俘获显微切割（LCM）中，通过使用整合到配备了连续成像系统的标准显微镜中的低能量红外激光，从标准的组织学切片中回收特定的细胞种群。该方法最初是由在NIH工作的激光技术小组开发的（Emmert-Buck et al. 1996，1997；Bonner et al. 1997），并通过与Arcturus公司（MountainView，California）的合作研究和开发协议而实现商品化。有关激光切割显微镜及其应用的信息，见公司的网站，关于附加方案以及有关LCM和激光切割显微镜的信息，可浏览Arcturus的网站www.arctur.com。

一旦组织经过正确地处理、切片、染色和脱水，就可以进行显微切割了（图5-4）。

一种名为 CapSure 的包被了热塑性薄膜的装置（小盖子）直接放置在准备好的组织切片的表面。低能量的近红外激光束促使薄膜与所选择的细胞接触，并将他们从周围的组织中提起。通过显微镜观察组织，采集初始路线图的图像以及切割前、切割后和盖子的图像，用于记录组织学、显微切割步骤和显微切割的细胞（图 5-5）。

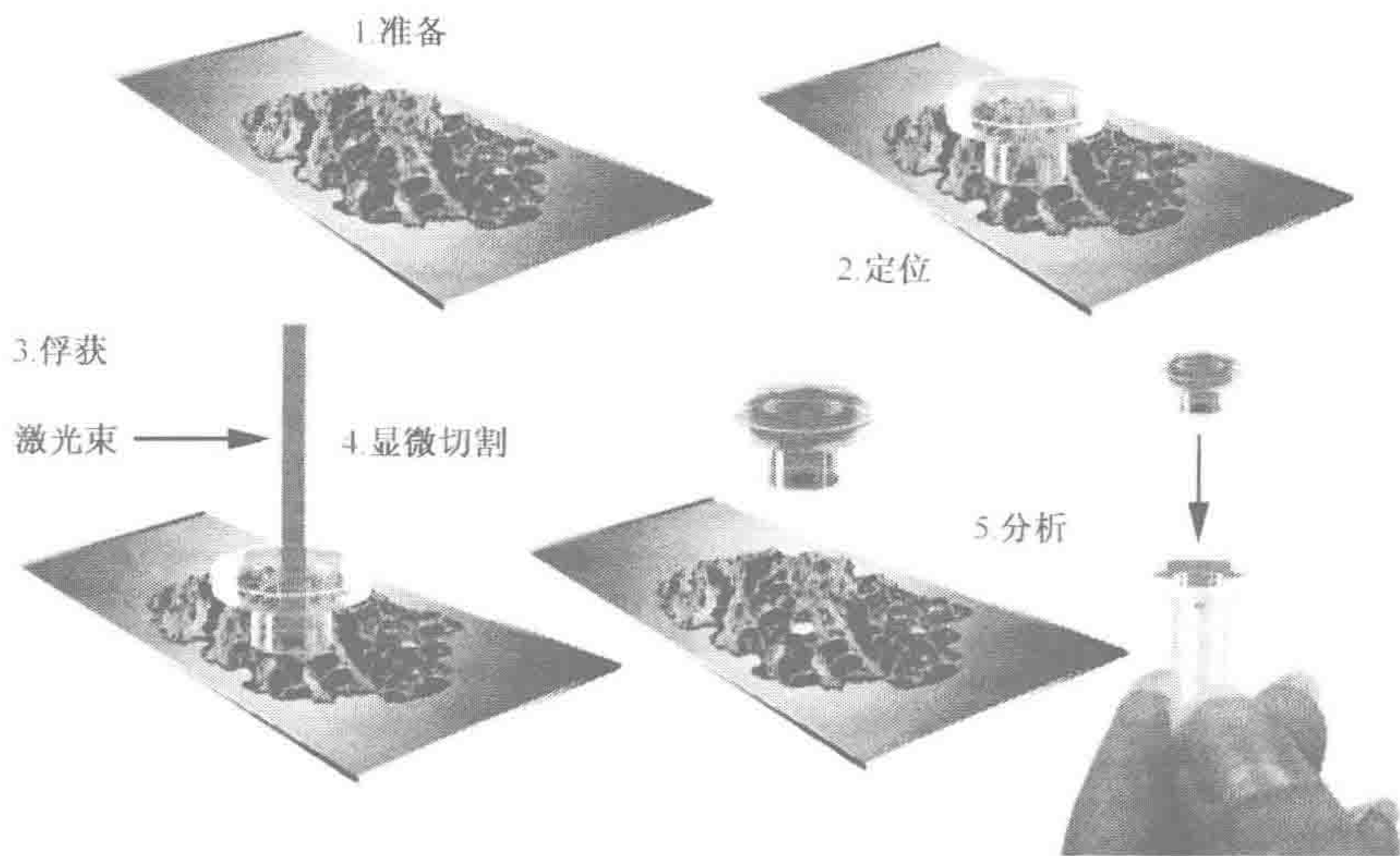


图 5-4　激光俘获显微切割

在载玻片上准备组织切片（1）并使用标准的倒置显微镜进行观察。对载玻片定位，使得所希望的细胞直接处在目标下方（2）。底部包被有转移膜的微量离心管盖放在细胞的上方，来自低能量近红外激光的短脉冲穿过管盖和薄膜打在细胞上（3）。膜发生融化，并与和它直接接触的细胞显微拓扑结构结合。盖子从组织上提起（4），所希望的细胞被分离出来用于后续的分子分析（5）

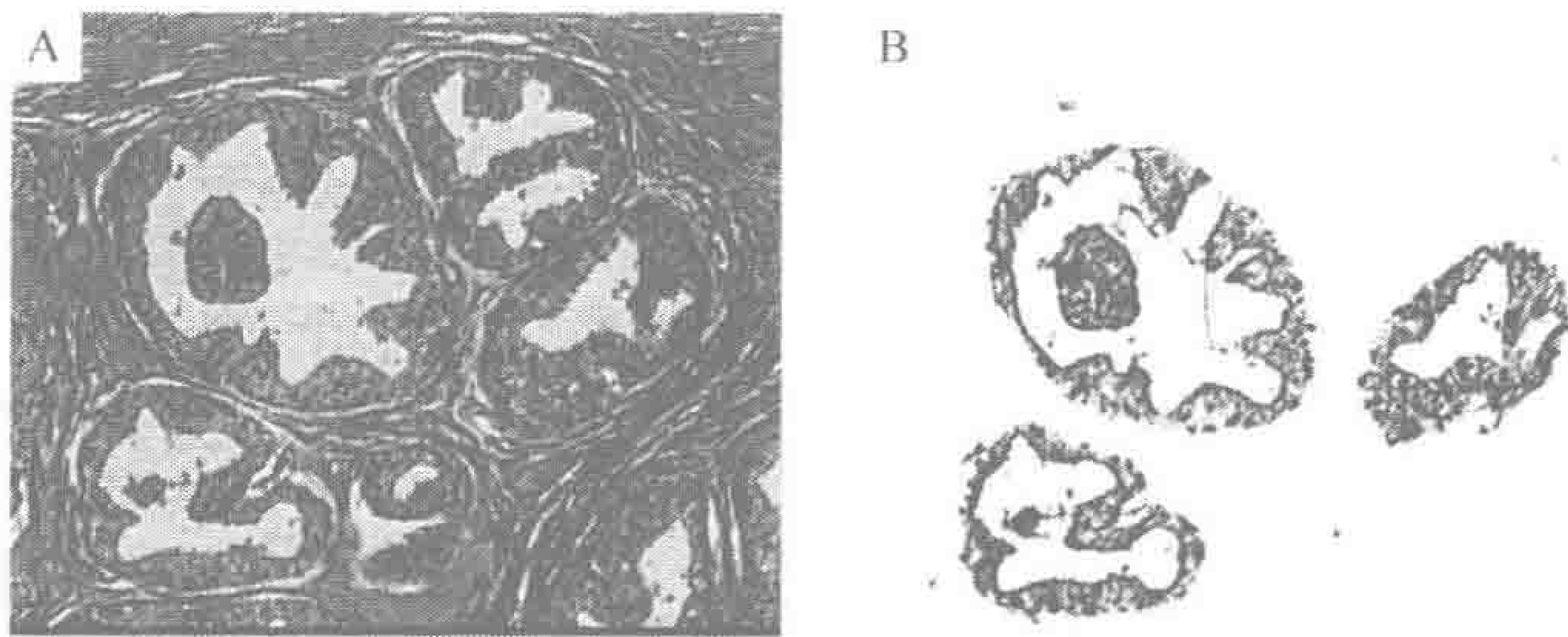


图 5-5　正常前列腺上皮的激光俘获显微切割

在显微切割前，切片用苏木精和曙红染色，并用二甲苯脱水。A. 正常的前列腺上皮是由数种类型细胞构成的复杂环境的一部分，激光使得仅仅上皮细胞附着到透明的热塑性薄膜上；B. 膜从组织切片上提起，取出所有被显微切割的细胞，并进行观察，显示对希望的组织进行了有效的俘获，而未被其他类型的细胞污染

当前，红外激光可以被聚焦到直径 7.5～30 μm；因此，应用标准的 LCM 系统去俘获单个细胞局限于那些尺寸较大或者分离较好的靶细胞（Luo et al. 1999）。但是，最近报道了一种改进的 LCM 显微镜，其中一个圆柱形的、40 μm 宽的薄膜表面以最小的接触作用力放置在样品上，并且在单个细胞被快速俘获和浓缩时进行转动（Suarez-Quian

et al. 1999)。此外，Arcturus 已经开发了一种高灵敏度、小体积的盖子用于 LCM 样品的俘获，这种盖子放在组织样品上方而不与之发生接触，激光使薄膜充分地熔化，以使它到达切片表面并与所希望的细胞发生结合。

322 材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

蛋白酶抑制剂混合片剂 Complete miniprotease inhibitor cocktail tablets(Roche)

对于所有的蛋白质分析，除了二甲苯以外，每 10 ml 各种试剂要溶解一片蛋白酶抑制剂混合片剂。

去离子水

Eosin Y 溶液（Sigma）

乙醇（70%，95%，无水的）

据报道，有些批次的乙醇（100%或 200 标准强度）中含有会使微阵列试验中产生很深荧光背景的污染物质，95%的乙醇不存在这个问题。为避免这个问题，可以用 95%的乙醇或已鉴定证明不会产生高荧光背景的无水乙醇。Warner-Graham 公司（Cockeysville，Maryland）出售的无水乙醇（USP 级乙醇，64-17-5）效果很好。

显微切割样品的抽提缓冲液（选择适合于特定研究的缓冲液，也就是说 DNA、RNA 或蛋白质分析：见步骤 20）

DNA 抽提缓冲液

50 mmol/L Tris-Cl(pH 7.6)

1 mmol/L EDTA(pH 8.0)

1% Tween-20

2 mg/ml 蛋白酶 K

RNA 抽提缓冲液

RNA 可以使用很多商品化试剂盒中的任何一种从显微切割的细胞中回收。通常，细胞在含有 β-巯基乙醇的标准的异硫氰酸胍的缓冲液中匀浆，然后进行 RNA 的纯化，如使用 QIAGEN的 RNeasy 试剂盒。在大多数情况下，当 RNA 仍然结合在硅胶膜上时，在清洗步骤之前使用 DNase Ⅰ消化以除去残留的基因组 DNA。更详情的信息，见第 3 章，以及本章的方案 1 或方案 4。

蛋白质抽提缓冲液

如果将蛋白质用于电泳分析，将细胞在下面的缓冲液中裂解：

8 mol/L 尿素

2 mol/L 硫尿

2%(*m/V*)CHAPS

1%(*m/V*)二硫苏糖醇（DTT）〈!〉

0.8% Pharmalyte 两性电解质载体（pH 3～10）

更多信息，见 Banks 等（1999），Emmert-Buck 等（2000）和 Ornstein 等（2000）。

如果从冰冻切片中回收和分析天然蛋白质，见方案 1 中**替代方案：冰冻组织样品的显微切割**的导言部分。

甘油（超纯级，GIBCO）

苏木精溶液（Mayer，Sigma）

二甲苯（混合）（ACS 试剂，Sigma）

细胞和组织样品

323

组织切片

从经甲醛固定、石蜡包埋的组织切下的切片裱贴在载玻片上，在室温或者低于室温下保存。直到显微切割之前，不要给切片脱蜡。包埋在低熔点聚酯中的切片在 4℃下保存。冰冻切片在－80℃或者更低的温度下保存。

专用设备

黏性小片（如 CapSure Cleanup Pad，Arcturus，Inc.）

GeneAmp PCR 管（0.5 ml 薄壁，不含 RNase）（Applied Biosystems）

建议使用这种试管，因为它们在盖上 LCM 盖子并将试管倒转时不会产生漏液现象。

激光俘获显微切割系统（PixCell Ⅱ LCM 仪器，Arcturus，Inc.）〈!〉

带有转移膜的 LCM 盖子（Arcturus，Inc.）

方法

载玻片的处理

如果要对石蜡包埋的切片进行染色，从步骤 1 开始本方案。对于冰冻包埋的切片，从冰箱中取出后稍作融化（如放在手背上）大约 30 s，然后从方案的步骤 4 开始。这个过程使得切片与载玻片结合更加紧密，并且切片表面更为“粗糙”，从而与 LCM 盖子的结合更佳。对于聚酯包埋的切片，将载玻片在无水乙醇中浸泡 2 次，每次 5 min（使切片脱蜡），然后从步骤 3 开始。

1. 为使组织切片脱蜡，将载玻片在新鲜的二甲苯中浸泡 5 min。
2. 将载玻片转移到新鲜的二甲苯中，浸泡 5 min。
3. 将载玻片转移到无水乙醇中，浸泡 30 s。
4. 将载玻片转移到 95%的乙醇中，浸泡 30 s。
5. 将载玻片转移到 70%的乙醇中，浸泡 30 s。
6. 将载玻片转移到去离子水中，浸泡 30 s。
7. 将载玻片转移到 Mayer 苏木精溶液中，浸泡 30 s。
8. 将载玻片转移到去离子水中 15 s；重复一次。
9. 将载玻片转移到 70%的乙醇中，浸泡 30 s。
10. 将载玻片转移到曙红 Y 中，浸泡 15 s。

11. 将载玻片转移到95%的乙醇中，浸泡 15 s。
12. 将载玻片转移到95%的乙醇中，浸泡 15 s。
13. 将载玻片转移到无水乙醇中，浸泡 15 s。
14. 将载玻片转移到无水乙醇中，浸泡 15 s。
15. 将载玻片转移到二甲苯中以确保切片脱水。载玻片浸泡 60 s。
16. 将载玻片在空气中干燥约 2 min 或者用空气枪轻轻处理使得二甲苯被完全除去。

    组织切片在 LCM 之前必须要完全干燥。使用 Accuduster 或者相似的设备帮助干燥以便有效的显微切割。

324 切割

为了保持 RNA 的完整性，必须在载玻片准备好之后立即进行显微切割，因为在室温下 1 h 之后，完全脱水的组织切片中 RNA 会发生显著的降解。染色后的载玻片可以在二甲苯中短期保存直到显微切割之前的干燥处理。在俘获完成后，要尽快使用变性缓冲液对细胞进行抽提。

17. 把仪器的盖子直接放在组织的表面。

    接近或者要切割区域的组织表面上没有不规则的形状是非常重要的。皱褶会使 LCM 盖子离开组织表面，从而在激光触发时减弱与薄膜之间的接触。此外，有时在组织表面上产生的微小不规则形貌（位于 LCM 盖子下方）无法观察到。但是，这类问题可以通过减少激光触发斑点的大小来发现。可以通过向盖子的支撑臂上增加额外的重物，或者临时性地增加激光的强度，从而部分或全部地消除这一问题。

18. 按照 PixCell 的说明对激光束进行聚焦。使用低能量红外激光传递一个脉冲。

    在对不同类型组织进行显微切割之前，应当根据经验来快速地确定合适的脉冲持续时间和激光功率的电流 mA。简要地说，脉冲越长，能量越高，俘获细胞的范围越大。对于所需要的斑点大小、功率和持续时间的设置可以通过在一块没有组织的区域上进行触发实验而优化。

    激光促使薄膜熔化，使其与组织相接触，所希望的细胞黏附到 LCM 盖子的薄膜上。激光束的大小可以在 7.5～30 μm 之间进行调节，以便于根据研究的目的去显微切割成群的或者是单个细胞。

19. 在显微切割完成之后，从载玻片上提起盖子和被显微切割下来的细胞，并使用一块黏性小片从 LCM 盖子上除去非特异性附着的细胞。把盖子分三次放在黏性小片上，在显微镜下观察以保证所有非特异性的细胞均被去除。

    Arcturus Engineering 开发了一种新型的盖子（HS LCM 盖子）能够消除同时带起不需要的细胞的现象。使用这些盖子，薄膜与组织之间保持 12 μm 的距离，并且只有目标细胞与薄膜发生接触。这个方法免除了在 LCM 中需要除去非特异性附着的材料。详细情况，见 www.arctur.com/technology/lcm-steps.htm.

20. 把盖子放在装有 100 μl 合适缓冲液的 0.5 ml GeneAmp 试管中用于分子分析（如用于 RNA 抽提的变性缓冲液）。倒转试管以使细胞裂解。

    当 ExtracSure 装置与 CapSure HS 联合使用时，可以使用 10 μl 缓冲液来抽提样品。

    本章方案 4 中给出了从显微切割的组织中抽提 RNA 的详细方法。Arcturus 开发了一

种称作 PicoPure RNA 分离试剂盒的 RNA 分离试剂盒（www.arctur.com），专门为LCM样品而设计。

方案 3

免疫-LCM

经过常规染色或者未染色冰冻组织切片的显微切割已经被成功地应用在获取纯化的细胞种群来进行基因表达分析上。然而，由于仅仅根据形态学来确定细胞类型和结构较困难，使得显微切割的精度和有效性常常受到限制。因此，研究了可以和 LCM 联用的快速免疫染色方法，并从冰冻切片中抽提出 RNA(Fend et al. 1999a)。通过免疫染色，能够分析从特定细胞种群所抽提的 mRNA，这些细胞种群通过其免疫表型或者是表达功能相关的抗原而进行鉴别。对于在丙酮、甲醇或者乙醇/丙酮中固定的切片，最佳的免疫染色要在 12～25 min 的总处理时间后才能得到。冰冻组织切片可以在无 RNase 的条件下，使用快速的三步链亲和素-生物素技术，然后脱水来进行免疫染色。经过免疫染色的切片就可以使用 LCM 进行切割（方案 2）。对于需要的（或不需要的）细胞类型鉴别能力的提高，使得显微切割的特异性、精度和速度都得到了显著的增加。从免疫染色的组织中回收的 mRNA 质量很高：使用相当于小于 500 个免疫染色和显微切割细胞的 cDNA，单步聚合酶链反应（PCR）就可以用于扩增长度超过 600 bp 的片段，不论是来自于持家基因（如 β-肌动蛋白），还是细胞特异性的信息（如 CD4 或 CD19）。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。
使用前将储存液稀释到合适的浓度。
丙酮
经 DEPC 处理的水（Research Genetics）
二氨基联苯胺（DAB）〈!〉
乙醇（70%，95%，无水的）
苏木精溶液（Mayer，Sigma）
磷酸盐缓冲液（PBS）（pH 7.4）
磷酸盐缓冲液，含 Ca^{2+} 和 Mg^{2+}
胎盘 RNase 抑制剂（Perkin Elmer，Branchburg）
二甲苯（混合）（ACS 级，Sigma）〈!〉

细胞和组织样品

组织块，速冻的

326 专用设备

冰冻切片机

DAKO Quick Staining 试剂盒（DAKO Corp）

该试剂盒中提供了用于三步链亲和素-生物素染色过程的材料，含有为短时染色而优化的预稀释单克隆或多克隆（兔）一抗。如果是由研究人员提供的一抗，其稀释比例要根据情况分别确定。在 DAKO Quick Staining 试剂盒中，二抗是以预稀释的形式来提供的。向一抗和 DAB 溶液中加入胎盘 RNase 抑制剂，浓度为 200～400 单位/ml。所有溶液均使用经 DEPC 处理的水配制。

载玻片（Superfrost Plus，Fisher Scientific）

方法

1. 在标准的冰冻切片机上，使用一片新的一次性冰冻切片机刀片切取速冻组织块，切片的厚度为 8 μm。
2. 将组织切片裱贴在 Superfrost Plus 载玻片上，并立刻将载玻片放入－80℃的冰箱中。
3. 在室温下将冰冻切片解冻 30～60 s。
4. 将样品迅速浸泡到冷的丙酮中 5 min 以固定样品。
5. 用 PBS 快速漂洗载玻片。
6. 使用 DAKO Quick Staining 试剂盒，将载玻片在室温下分别和一抗、二抗以及第三种试剂温育，每次 90～120 s，对组织切片进行免疫染色，每步之间使用 PBS 短时漂洗。
7. 用 DAB 显色 3～5 min，并用苏木精复染 15～30 s。
8. 将切片依次浸泡在一系列 70％、95％、无水乙醇中脱水，每次 15 s。
9. 将切片在二甲苯中浸泡 2 min，重复一次。
10. 将样品在空气中晾干。

免疫染色切片现在可以按照方案 2 中所述进行 LCM。

327 方案 4

应用于表达分析的显微切割组织的处理

下面的方案中介绍了使用 Stratagene，Inc. 的 Micro RNA 分离试剂盒从显微切割的样品中回收 RNA（Ohyama et al. 2000）。其他一些方法也被证明能够成功地从 LCM 或者 LPC 样品中分离 RNA，如 Luo 等（1999）和 Bernsen 等（1998）。同时见本方案最后**用于从少量显微切割的细胞分离 RNA 的商品化试剂盒**信息栏。从显微切割的样品中回收的 RNA 可以按照本指南第 3 章中 PART Ⅱ 的方案所述进行标记和与微阵列杂交。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录3。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录2。

使用前将储存液稀释到合适的浓度。

β-巯基乙醇（14.2 mol/L）〈!〉

氯仿：异戊醇（24：1）〈!〉

糖原的水溶液（10 mg/ml）

异丙醇（冰冷）〈!〉

酚（水饱和的，pH 5.2）〈!〉

RiboGreen RNA 定量试剂（Molecular Probes，Inc）

RNA 变性缓冲液（GITC）

  4 mol/L 异硫氰酸胍（GITC）

  0.02 mol/L 柠檬酸钠

  0.5%十二烷基肌氨酸钠

醋酸钠（2mol/L）用乙酸将 pH 调节为 4.0〈!〉

酶和缓冲液

DNase Ⅰ(10单位/μl)(GenHunter)

10×DNase 缓冲液

RNase 的蛋白质抑制剂（如 Promega Inc. 的 RNasin 或者 5 Prime→3 Prime 的 Prime Inhibitor 引物抑制剂）

细胞和组织样品

位于 LCM 盖子上的显微切割组织（来自方案2）

专用设备

Bioanlyzer（Agilent Technologies）

  可选择的，见步骤27。

GeneAmp PCR 管（0.5 ml 薄壁，不含 RNase）（Applied Biosystems）

  建议使用这种试管，因为它们在盖上 LCM 盖子并将试管倒转置时不会产生漏液现象。

荧光光度计（需要少量样品的，如 Turner Designs 的 TD-60）

328

方法

重要事项：在开始本方案之前，用可以使 RNase 失活的产品清洗所有的移液管（如 RNase*Zap*

Wipes，Ambion Inc.）。

RNA 的抽提

1. 为每个盖子上要处理的组织准备一支 GeneAmp 微量离心（0.5 ml）薄壁 PCR 管，其中装有 200 μl 的 RNA 变性缓冲液和 1.6 μl 的 β-巯基乙醇。

   关于从显微切割材料中抽提 RNA 的替代方案，见方案 8。

2. 将带有显微切割组织的 LCM 盖子放到步骤 1 中所准备的 GeneAmp 试管上。将盖好盖子的试管在 2 min 内翻转数次，以便从盖子上回收组织。

   对于石蜡包埋的标本，在 60℃下温育 20 min（在试管翻转时），有助于后面 RNA 从组织中释放出来。

3. 将溶液从 PCR 管中取出并放在一个牢固的、不含 RNase 的 1.5 ml 微量离心管中。
4. 向样品中加入 20 μl（容量的 1/10）2 mol/L 醋酸钠（pH 4.0）。
5. 加入 220 μl（等体积的）水饱和酚。
6. 加入 60 μl（酚体积的 1/3）氯仿：异戊醇。
7. 将试管用力振摇 15 s。
8. 将试管在融化的冰上放置 15 min。
9. 将试管在微量离心机中于 4℃下离心 30 min 使得水相和有机相分离。

   相分离也可以通过使用预离心的锁相胶试管（Phase Lock Gel 5 Prime→3 Prime）来进行。

RNA 的回收

10. 将上层水相转移到一支新的试管中。

    任何偶然转移入水相的下层有机相都会影响到后续的异丙醇沉淀。为了从水层中除去残留的有机相，加入一倍体积的 100% 的氯仿，混合均匀，然后在 4℃下离心 10 min 使得水相和有机相分离。将上层转移到一支新的试管中。

11. 向水层中加入 1～2 μl 糖原（10 mg/ml）以及 200～300 μl（相同体积）冷的异丙醇。

    糖原有助于观察离心得到的团块，通常在使用少量的 RNA 时是很难看见的。

12. 将样品在−80℃下放置至少 30 min（或者过夜）。

    开始离心之前，如果试管在异丙醇沉淀时冻结，则需要轻度的解冻。

13. 在微量离心机中，于 4℃下以最大速度离心 30 min 来回收 RNA，盖子的铰接部分朝向外侧，这样可以较好的预测团块的位置。
14. 弃去上清液，用 300 μl 冷的 70% 乙醇清洗团块，然后在 4℃下离心 5 min。
15. 弃去上清液。
16. 将团块在冰上晾干而除去残留的乙醇。

    过度干燥会使得团块不容易被重悬于溶液中。

17. 将团块保存在−80℃下直到使用，或者按照下面步骤中所述用 DNase Ⅰ进行处理。

受到基因组 DNA 污染的 RNA 制备物会带来问题，很可能是由于在组织处理过程中产生的小的 DNA 片段。

使用 DNase 处理 RNA

18. 向每个 RNA 团块中加入 15 μl 经过 DEPC 处理的水和 1 μl(20 单位) RNasin 或者其他 RNase 的蛋白质抑制剂。轻弹每支试管的侧壁使得组分混合直到团块被溶解。
19. 将样品在微量离心机中短暂离心。
20. 向每个样品加入 2 μl 的 10×DNase 缓冲液和 20 单位（2 μl）的 DNase Ⅰ。将试管在 37℃下温育 2 h。

    一些研究人员省略了 DNase Ⅰ处理，因为这可能需要另一轮的 RNA 纯化并伴随着珍贵材料的损失。他们认为低水平的基因组 DNA 污染未必会影响到转录样式，因为从基因组 DNA 的反转录是稀少的事件，而且 DNA 微阵列上各点发出的信号中源自 DNA 的信号很容易鉴别出来，因为它们相对于细胞类型而言是中性的。

    如果按照本方案所述进行 DNase 处理导致 RNA 的明显丢失，更好的方法是使用离心柱-树脂（spin-matrix method）方法对 RNA 进行纯化。这可以使得 RNA 仍然结合在树脂上时进行 DNase 消化（见本方案最后**用于从少量显微切割的细胞分离 RNA 的商品化试剂盒**信息栏）。

21. 向每个样品中加入下列试剂抽提 RNA：

2 mol/L 醋酸钠（pH 4.0）	2 μl
酚（水饱和的）	22 μl
氯仿：异戊醇	6 μl

22. 将试管用力振摇 15 s，然后在融化的冰上放置 5 min。
23. 将试管用微量离心机在 4℃下以最大速度离心 10 min。

330

24. 将上层转移到一支新试管中。

    任何偶然转移入水相的下层有机相都会影响到后续的异丙醇沉淀。为了从水层中除去残留的有机相，加入一倍体积的 100%的氯仿，混合均匀，然后在 4℃下离心 10 min 使得水相和有机相分离。将上层转移到一支新的试管中。

25. 重复上面的步骤 11～15 继续 RNA 的抽提，对异丙醇的体积作相应的调整。
26. 晾干团块，但注意不要过分干燥。将团块用 19 μl 经过 DEPC 处理的水和 1 μl RNase 抑制剂重悬，并于−80℃下保存。
27. 如果有可能，检查 RNA 的质量。

    在此之前，评价这些样品的质量是非常困难的，因为从 LCM 回收的 RNA 的量是如此的少，并且使用标准方法检测质量带来的损失是不具备充分的优势的。然而，最近出现的 Bioanalyzer（来自 Agilent Technologies），由于能够评价少至大约 5 ng 的 RNA 的质量，这使得在任何后续的分析开始之前检查 RNA 的质量逐渐成为可能。

28. RNA 的定量。

    从显微切割的样本中回收的 RNA 的量太少而无法使用普通的光度计进行测量。作为替代，RNA 的定量使用，如 RiboGreen RNA 定量试剂（Molecular Probes Inc.）和一

种需要少量样品的荧光光谱仪（如 Turner Designs 的 TD-60）。注意像 RiboGreen 这样的试剂与 DNA 和 RNA 均可以结合。因此，RNA 定量应该使用经过 DNase 处理的样品。

现在可以对样品进行后续的处理，而且可以按照第 3 章的 PART Ⅱ中的方案之一进行标记用于微阵列分析。

用于从少量显微切割的细胞分离 RNA 的商品化试剂盒

一些商品化的试剂盒使得在通过 LCM 从组织切片回收的少量细胞中来分离总 RNA 这一任务得到大大的简化。最佳结果来自于使用离心-树脂柱的试剂盒，如下面介绍的。

- Absolute RNA Kit(Stratagene)（http://www.Stratagene.com/cellbio/purification/absolutely_rna_micro.htm）. 用 LCM 回收的细胞被裂解，与乙醇混合，并加到基于硅胶的纤维树脂上。开始洗脱步骤之前，当 RNA 仍然结合在硅胶树脂上时，用 DNase I 消化除去残留的基因组 DNA，然后对柱子进行洗脱，流出较小体积的 RNA。
- PicoPure Kit(Arcturus，Inc.)（http://www.arcturus.com/products/picopure_rna_isolation.htm）。该试剂盒与 Stratagene 的 Absolute RNA Kit 非常类似，不同之处在于 DNase Ⅰ消化在 RNA 从柱子上洗脱下来之后进行。该试剂盒是为从 LCM 分离得到的 10～1000 或者更多的细胞中回收 RNA 而优化的。
- Micro RNA Isolation Kit(Stratagene) 通过在基于胍/β-巯基乙醇的缓冲液中裂解细胞，然后用酚：氯仿抽提，并在糖原存在下用乙醇沉淀来纯化总 RNA。RNA 重新溶解后，对它进行 DNase Ⅰ处理以除去污染的基因组 DNA。该方法与方案 4 中上面介绍的方法非常相似。
- Rneasy Minikit(QIAGEN)（http://www.qiagen.com/literature/brochures/pcr/pdf/pcrcha28.pdf）。该微型试剂盒主要用于如第 3 章中方案 1 所介绍的。详细情况见 Ohyama 等（2000）。

331 PART Ⅱ 激光压力弹射

Karin Schütze[a], Bernd Becker[b], Monique Bernsen[c], Tone Björnsen[d], Dieter Brocksch[e], Christer Bush[d], Annette Clement-Sengewald[f], Marcory C. R. F van Dijk[c], Gabriele Friedemann[a], Wolfgang Heckl[i], Georgia Lahr[j], Per Lindahl[k], Annette Mayer[l], Sven Nilsson[k], Stefan J. Scheidl[k], Monika Stich[j], Wilhelm Stolz[b], minoru Takemoto[k], Stefan Thalhammer[i], Thomas Vogt[b], Renate Burgemeister[a]

（a. P. A. L. M. 微激光技术公司，贝恩里德市，德国；b. 罗根堡大学，罗根堡，德国；c. 奈梅亨大学医学院，奈梅亨，荷兰；d. University Hospital，Department of Pathology，Regionsykehuset Tromso-Rito，Tromsoe，Norway；e. 卡尔蔡斯显微镜集团公司，哥廷根，德国；f. 慕尼黑大学，慕尼黑，德国；i. 晶体学研究所，慕尼黑大学，慕尼黑，德国；j. 慕尼黑-哈拉兴大学医学院，慕尼黑，德国；k. 哥德堡大学，医学生物化学系，哥德堡，瑞典；l. 人类学和人类遗传学研究所，哥德大学，法兰克福，德国）

导言

激光压力弹射（LPC）是由 P. A. L. M. GmbH(www.PALM-Microlaser.com）的开创研究人员研制的一种用于分离感兴趣的细胞群或者单个细胞的方法。LPC 使用激光诱导的光分解作用产生的气体压力将单个细胞或部分组织从平面上提升起来，并弹射

（高达数厘米）到装有缓冲液的容器中，通常是常规的微量离心管管盖或是一块微量滴定板，放置在激光束的光路中（图 5-6）（Schütze and Lahr 1998；Lahr et al. 2000；Bertheau et al. 2001；Westphal et al. 2002）。俘获的标本经过离心集中到微量离心管的尖端部分。根据组织的特性和标本制备的方法，可以先用激光圈出单个细胞或者是组织切片上一个选定的区域，在靶细胞和不需要的相邻细胞之间形成分隔。位于所希望区域中不需要的材料可以选择性地清除，从而通过一个被称作激光显微切割和压力弹射（LMPC，图 5-7）的过程提供了俘获纯粹样品的方法（Buerger et al. 2000a；Lahr et al. 2000；Stoehr et al. 2000；Vona et al. 2000；Hoffmann et al. 2001；Specht et al. 2001；Fink et al. 2002；Scheidl et al. 2002；Westphal et al. 2002）。

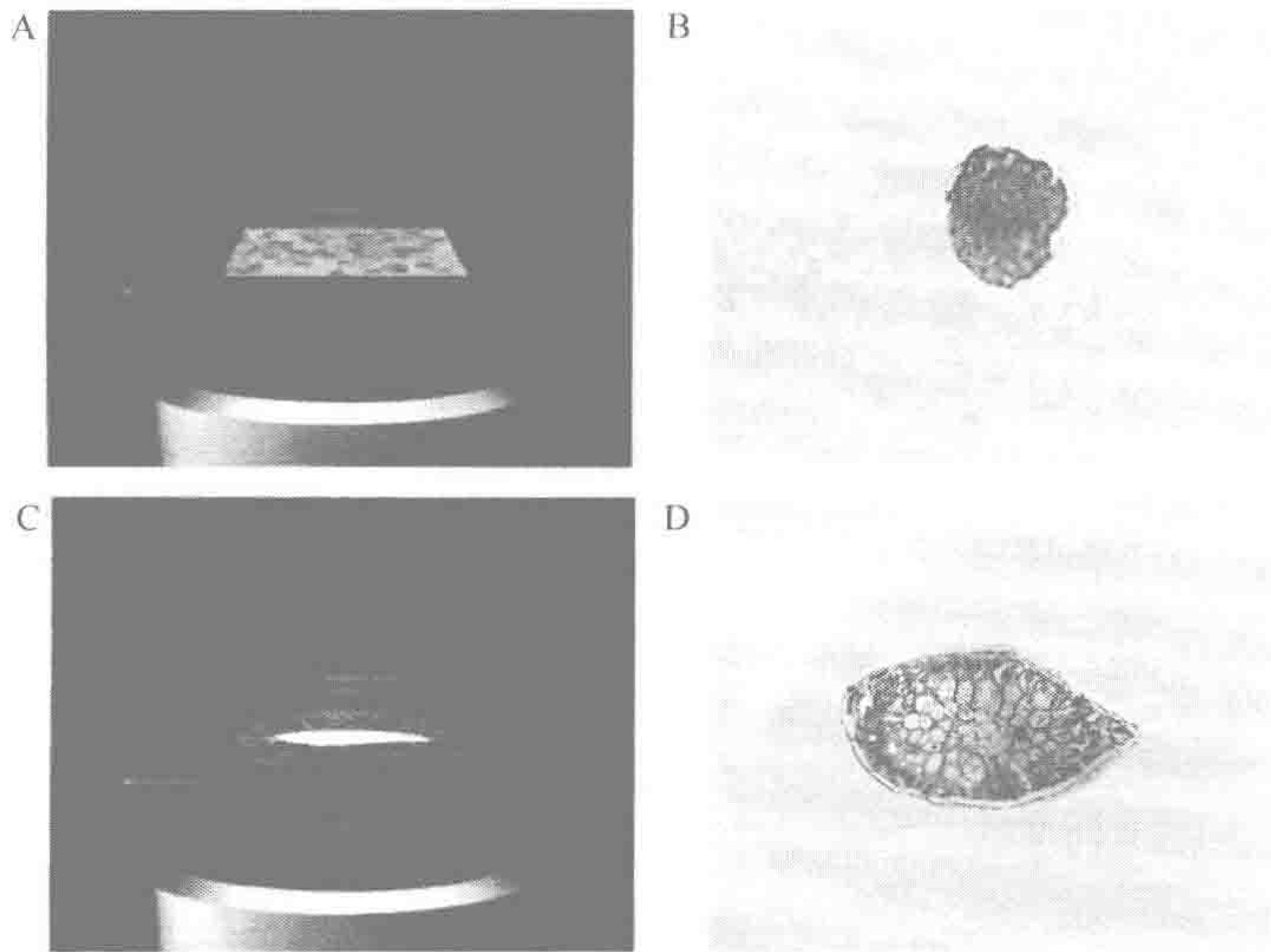


图 5-6　激光压力弹射（LPC）示意图

弹射自细胞离心涂片标本的单个细胞（A），并在俘获盖子上进行观察（B）。从裱贴在膜上的组织切片上选取一个区域进行显微切割和弹射（C），并在俘获盖子上进行观察（D）

332

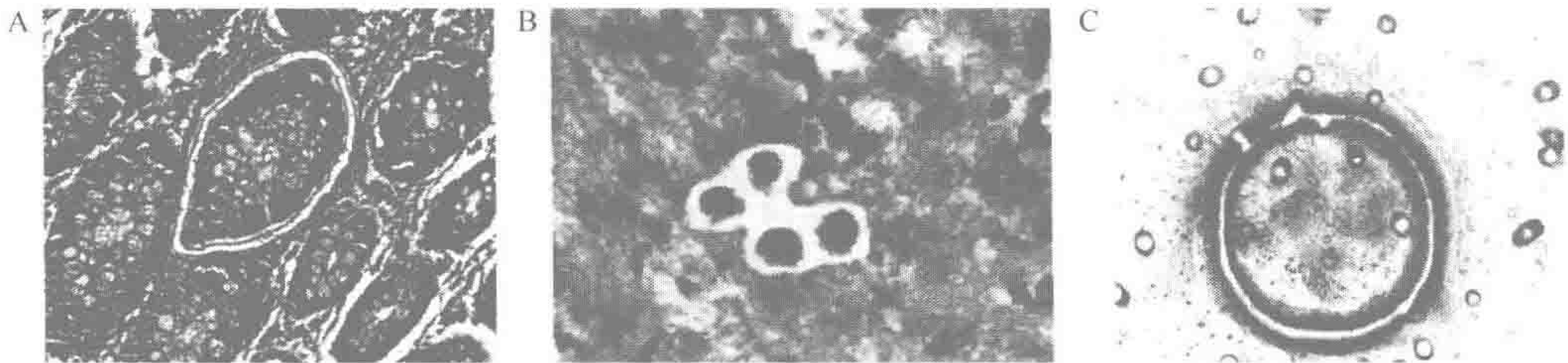


图 5-7　激光显微切割和显微手术示例

A. 在裱贴于膜上的组织切片上圈定一个大腺的范围（使用 20×物镜进行观察）；B. 对单个肿瘤细胞进行显微切割；C. 对滞留在常规血液过滤器（Teflon；20 μm 厚）中的循环肿瘤细胞进行显微切割，用于后续的遗传学评估（使用 40×物镜进行观察）

激光切割的尺寸取决于波长和光束的质量、物镜的放大倍数和数值孔径，以及所选择的标本的吸收特性。为了进行精确的聚焦，需要有高质量光束的激光器和数值孔径 N. A. >1 的物镜。使用 100 倍的油浸物镜（N. A. >1. 3）能够精确地分离单个细胞核、纤维、染色体，甚至部分的染色体。在激光有限的焦斑中，产生大于 1 MW/cm^2 的功率

密度，足以破坏生物物质的化学键。在激光切割过程中，物质被光分解为原子和小分子，并以超音速被吹走。因为这是一个光化学过程而且速度极快（亚微秒），伴随的对相邻物质的破坏可以忽略（Srinivasan 1986）。

此外，由于感兴趣的细胞可以在没有机械接触的情形下从切片上弹射开去，因此未选择的组织不可能造成污染。这使得在同一个收集设备上可以从切片的不同位置收集多个单元或者是对多张切片（如连续切片）进行回收。LPC 或 LMPC 是完全自动化的，可以适用于任何细胞或者组织制备技术与染色方法，同时可以与荧光照明联用。使用计算机图形学工具或者进行图像分析，研究人员对感兴趣的细胞或区域进行选择，勾勒出轮廓，并用彩色编码（图 5-8）。勾勒出的样品汇总在一个单元列表中，然后进行自动处理（表 5-2）。

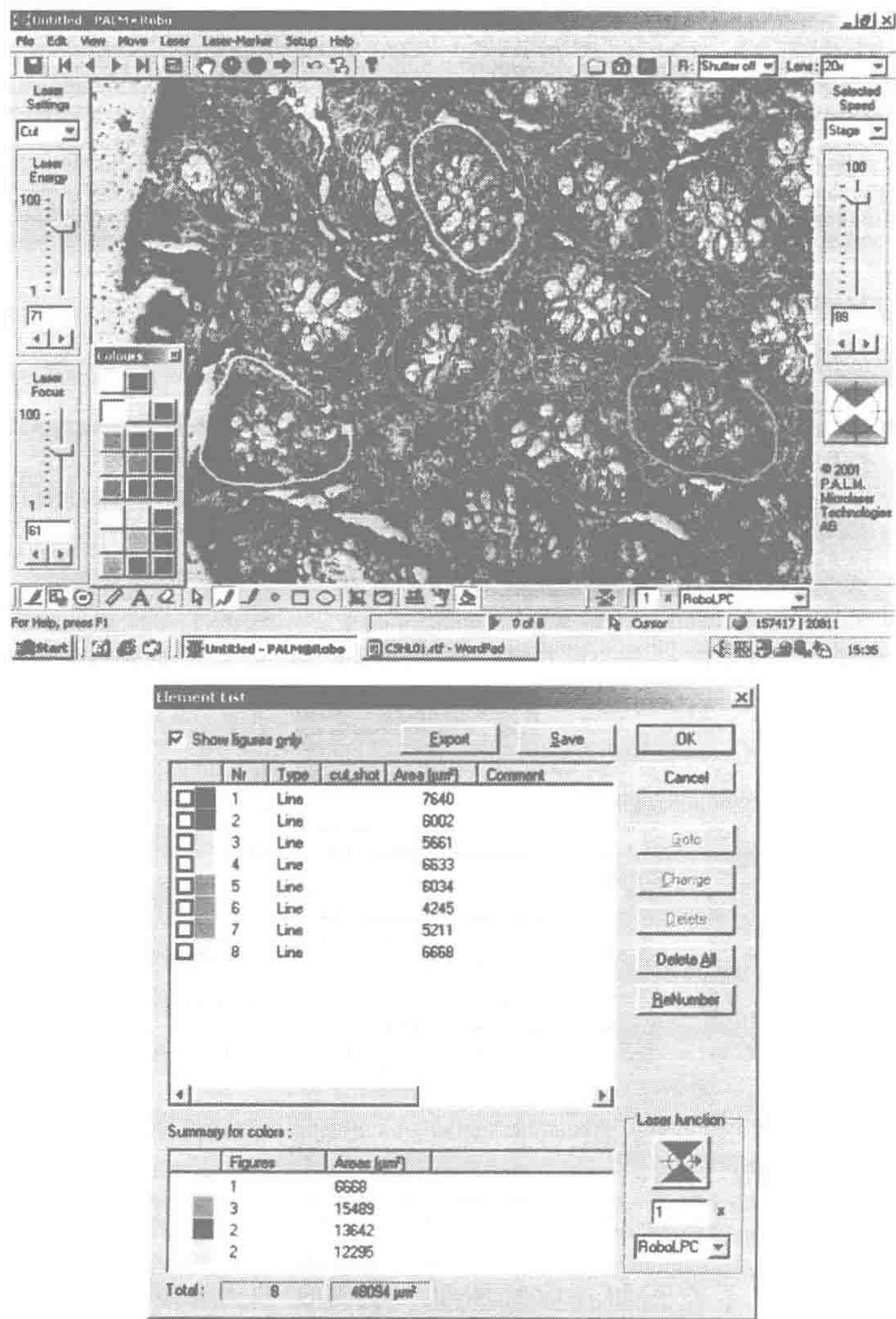


图 5-8　带有彩色编码的预选和自动显微切割

（上图）被选择样品带有色彩的轮廓和 RoboLPC 线条的显示，红紫色的点显示了弹射的位置。细微的光束自动地勾勒出标本的轮廓，并立即将其弹射到俘获瓶中。（下图）单元列表能够根据带有色彩编码的预选对样品进行分组；只有勾勒出轮廓的区域应用激光处理而被俘获

表 5-2　激光的动作以及自动显微切割和弹射方法

334

方法	图像	说明
Cut		激光沿着预先确定的路线切割
CloseCut		打开的图像被关闭，随后进行切割
LPC		只有用点标记的标本被弹射
AutoLPC(仅适用于裱贴在玻片上的标本)		对勾勒出的区域进行弹射，进行了多次激光触发(LPC 样式在激光"Setup 菜单"里预先选择并显示在显示器上)(图 5-9)
CloseCut&AutuLPC		打开的图像被关闭，且勾勒出的区域在 AutoLPC 之前被切割，以避免被相邻的组织污染
RoboLPC		勾勒出的标本被切割至仅剩下一个小的连接部分，在该部位整个区域通过一次激光作用被弹射(连接部位的大小在激光"Setup 菜单"里预先选择并与 LPC 点一齐显示)

仪器

335

PALM 微激光束、电动显微镜载物台以及自动显微切割和弹射的软件（包括现有的关于将 PALM 系统应用于日益增加用途的说明书）的详细情况在 PALM 提供的文献和公司的网站 www.PALM-Microlaser.com 中有介绍。通过软件（RoboSoftware）控制马达驱动的激光能量和聚焦设置，以及电动载物台的运动。仪器的设置在不同的用途以及特定的 LMPC 显微镜光学系统中各不相同。

弹射后的标本收集在放置于物镜上方和激光束之中的常用微量离心管盖子或者是其他任何合适的容器（即微量微定板）中。为了获得最佳观察效果，建议使用带有水平的未磨砂透明表面的盖子。使用带有狭窄边缘（<1 mm）的平顶盖可以提高激光俘获的效率，并使得5×、10×和长工作距离的40×物镜可以聚焦于盖子表面。为了使弹射的样品在收集管盖子中紧紧地附着，盖内要加入少许缓冲液（如RNA裂解缓冲液），这将使得弹射的标本易于通过快速离心步骤转移至试管中。作为替代，标本也可以弹射入充填了半黏性硅酮的平顶8排盖中。

标本的处理

用于LPC的标本可以使用分子病理学中常用的任何一种技术来处理。它们可以在液氮中冰冻并在−80℃下保存；可以在甲醛中固定，使用石蜡包埋，并用任何标准的荧光、免疫组织学、或者显色的组织学染色方法进行染色；也可以在OCT（最佳切片温度包埋复合物或Tissue-Tek）中包埋和保存。

- 在各种固定试剂中，沉淀试剂如甲醇、乙醇和丙酮，对于保持RNA的完整性和回收效果较好，但是它们不适合于冰冻切片。
- 交联试剂，如多聚甲醛和甲醛固定组织，收获的RNA量要少于速冻的或乙醇固定的组织。
- 石蜡包埋切片的形态学较好，但是与包埋在OCT中的冰冻样品相比，RNA的回收显著变差。
- 从使用OCT包埋的材料切下的切片经常可见微细的裂缝，但是保持了细胞间的相对位置。OCT与大多数包埋复合物（包括石蜡）一样，会干扰激光的切割，在LPC或LMPC之前必须从切片中除去（见方案7）。

裱贴在玻片上的标本

组织切片可以裱贴在常用的玻璃载玻片上，或者是经过等离子处理以改善组织和细胞附着的带（正）电的载玻片上（见方案5）。在通常的玻璃载玻片上的标本（例如，从细胞离心分离得到的单个细胞；图5-9A～图5-9C）容易被聚焦在位置略低于目的细胞的单次激光作用而弹射开去。

较大标本或者整个组织区域的弹射需要多次激光作用，因为每次激光作用只能得到一小部分。从玻璃载玻片上弹射通常在取样的盖子内形成细胞碎片或者组织薄片。但是，组织标本的碎裂并不会危及随后的分子分析（图5-9D～图5-9F）（Lahr 2000；Lahr et al. 2000；Bertheau et al. 2001；Scheidl et al. 2002；Westpha et al. 2002）。

裱贴在膜上的标本

在一些应用中，标本可以裱贴在特殊的薄膜上，这里薄膜充当了支架，在弹射过程中支持所希望的样品，并保持它的形态（图5-10）。

在激光显微切割过程中，选取的组织和下方的薄膜被激光束同时勾勒出来，并被单次激光作用而弹射。建议在下列情况采用裱贴在膜上的方法。

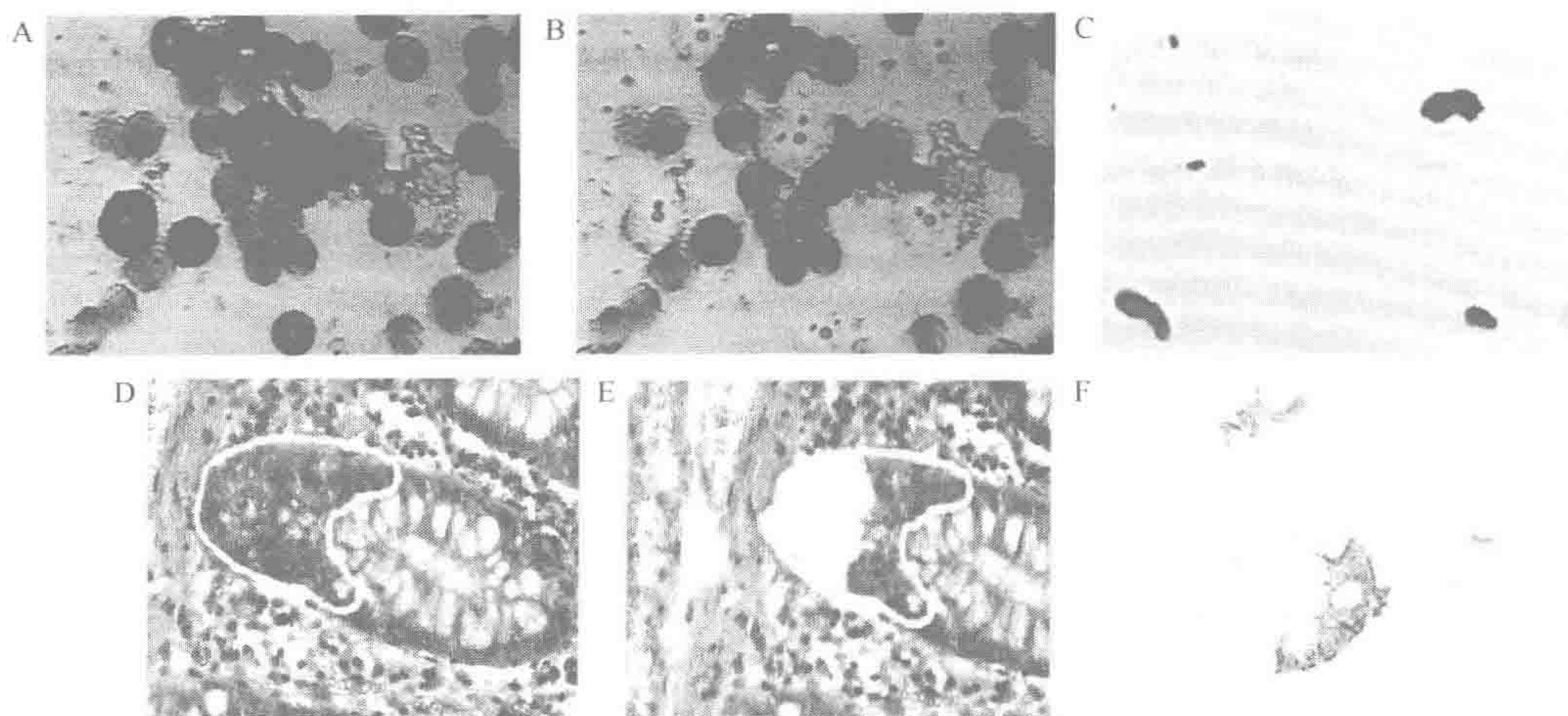


图 5-9　从细胞离心分离的标本上直接弹射单个细胞和组织区域

336

A. 用计算机图形工具标记细胞；B. 每个标记的细胞被自动重新定位，并通过单次激光作用使它们分别弹出，在载玻片上留下相应的缺口；C. 弹射的细胞在到达通常的微量离心管盖子中时，有时会碎裂。这不会影响到 DNA、RNA 和蛋白质的回收；D. 裱贴在载玻片上的组织切片的显微切割和弹射：弹射之前的显微切割在已选取的和未被选取的材料之间形成一个切割清晰的间隙，用于提供纯样品制备；E. 经过多次激光作用，选取的标本被一片片地弹射出去，并在载玻片上留下了相应的空点；F. 经过多次激光作用的弹射，常常在收集盖中形成组织薄片；然而，由于样品是来自于形态确定的区域，因此 DNA、RNA 或者蛋白质的回收不会因此而受到影响

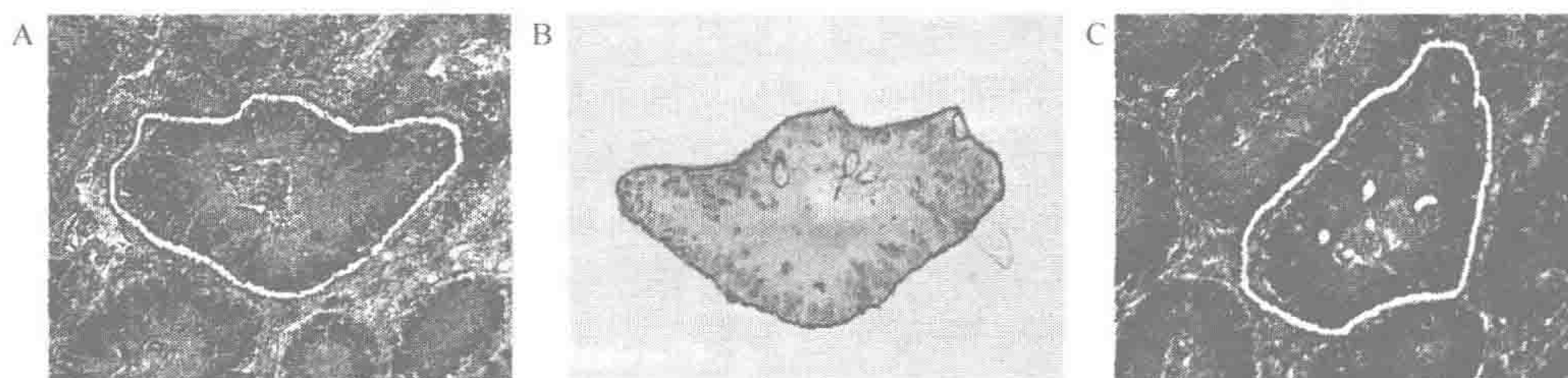


图 5-10　裱贴在膜上的标本的显微切割和弹射

A. 激光在组织学的组织切片中勾勒出选定的区域，同时切割下方的薄膜（使用 40× 物镜观察）；B. 弹射的样品进入俘获盖子中，并保持了形态学的完整性（使用 40× 长工作距离物镜观察）；C. 在弹射之前，显微切割和选择性地消除不需要的材料（使用 40× 物镜观察）

337

- 易碎的标本和与下层基底紧密相连（如伸展的单个细胞、血涂片，或者细胞涂片）的材料（Burgemeister et al. 1999）、染色体或者纤维。
- 组织切片上大的细胞范围。用这种方法，可在数秒内弹射直径高达 1 mm 的区域，相当于数千个细胞。
- 可以在膜上培养的细胞类型。这些细胞可以在显微切割和弹射之前进行固定，或者是在存活状态下被弹射和收集。细胞经过显微切割和弹射之后可以很好地存活，并且能够用作后续的培养和克隆（Mayer et al. 2002）。

第 5 章中这一部分的方案介绍了应用于组织切片的各种类型载玻片（方案 5），用

于切片的组织的处理方法（方案6～方案8）和使用“液体盖玻片”保护组织切片（方案9）。方案10中给出了用于俘获活细胞的最后一个方案。弹射所使用的仪器和方法的详细情况，见PALM网站www.PALM-Microlaser.com。网站中还有关于PALM MicroBeam各种应用所发表文章的更新列表。

338

方案5

准备载玻片用于为LMPC进行组织切片

激光显微切割和压力弹射可以处理常规固定和石蜡包埋的组织（方案6），或者是新鲜的或冰冻标本（方案7）。根据组织的特性，为了增加组织黏附，可能需要预先包被载玻片或者使用PALM MembraneSlides（在下面的方案中介绍）。

有两种类型的薄膜载玻片，由于它们的化学成分不同，激光对这些支持膜的切割行为也存在差异。要根据应用来决定使用某种特定类型的载玻片——任何一种均可能被采用。

- **PEN薄膜**。PEN（polyethylene naphthalate，聚萘二甲酸乙二醇酯）薄膜（可从PALM获取）厚度为1.35 μm，对UV-A波段有强烈吸收，易于激光切割。PEN薄膜主要用于常规的组织制备，如显微切割单个细胞或者大范围的细胞。但是，它也可以应用于从细胞涂片或者是在膜上培养的细胞中进行细胞回收。
- **POL薄膜**。POL（聚酯）薄膜厚度为0.9～1.0 μm，并且不如PEN薄膜那样对紫外线敏感。在弹射之前，重要的邻近材料或者部分标本本身必须去除以避免不希望的材料的污染，POL薄膜主要应用于这些场合下的制备。在不希望的组织区域被清除以后，选取的标本用强度稍微高一些的激光勾勒出轮廓而环切，产生小的“薄膜-标本岛状物”，随后被弹射到收集盖子中去。

PALM研制了一种特殊的“双层薄膜”载玻片，包含了一层厚的Teflon薄膜——与玻璃载玻片相似——和一层紧贴于其上的薄膜（POL或者PEN中的任何一种）。这些PALM DuplexSlides可以耐受化学处理，并使得LMPC可以应用高数值孔径但工作距离很短的物镜。在一些特殊的情况下，一种“薄膜-组织-薄膜”夹心结构有助于防止未黏附的或者是活的标本发生漂移。在这种情况下，一种特殊的MembraneFrame框架用于放置在薄膜载玻片的上方。在显微切割过程中，PALM激光的功率足够强烈，可以同时切穿两层薄膜以及夹在它们中间的标本。

为了进行活细胞培养和随后的弹射（见方案10），还研制了特殊的PALM ChamberSlides（附着了小烧瓶的载玻片）以及PALM DuplexDishes。

339

材料

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录2。

使用前将储存液稀释到合适的浓度。

聚-L-赖氨酸溶液（0.1%，*m/V*）（P 8920，Sigma）

专用设备

Superfrost Plus 显微镜玻璃载玻片（6166 11 55，Fisher Scientific）或者 PALM MembraneSlides

方法

1. 为了包被玻璃载玻片，将一滴聚 L-赖氨酸溶液加到载玻片上。
2. 使用另一张载玻片或者是吸头将溶液均匀分布在载玻片上。
3. 将载玻片在室温下干燥 30 min。
4. 将用于切割的组织标本放到玻璃的或者是包被了膜的载玻片上（见实验 6、7 或 8）。

方案 6

340

用于 LMPC 的石蜡包埋组织切片的制备

一般来说，如果要从俘获的材料中抽提 RNA，切片应在尽可能避免暴露于 RNase 的条件下进行切片和处理。所有的仪器和染色缸都必须用去除 RNase 污染的试剂，如 RNaseZap 处理，并且所有溶液均需用不含 RNase 的成分新鲜配制。更多细节，见第 3 章中 **RNA 操作的特别注意事项**的信息栏。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

重要事项：本方案中使用的所有试剂用经过 DEPC 处理的 H_2O 配制（见第 3 章中 **RNA 操作的特别注意事项**的信息栏）。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

乙醇（无水的，96%，70%）

据报道，有些批次的乙醇（100%或 200 标准强度）中含有会使微阵列试验中产生很深荧光背景的污染物质，95%的乙醇不存在这个问题。为避免这个问题，可以用 95%的乙醇或者已鉴定证明不会产生高荧光背景的无水乙醇。Warner-Graham 公司（Cockeysville，Maryland）出售的无水乙醇（USP 级乙醇，64-17-5）效果很好。

RNase*Zap*（9780，Ambion）

染色液：Mayer 苏木精，甲基绿，或甲苯胺蓝二甲苯〈!〉

建议在微阵列实验之前测试特殊染色液的影响；初步的结果表明苏木精-曙红与“细胞核快

红”染色一样，都可以得到最好的结果。

细胞和组织

组织标本（石蜡包埋的）

专用设备

玻璃或者膜包被的载玻片（见方案 5）
显微镜用薄片切片机

341

方法

1. 选择常规的石蜡包埋标本，并用标准的薄片切片机切成 3～15 μm 厚的薄片。
2. 将切片直接转移至玻璃载玻片或者薄膜载玻片上。
3. 将载玻片在 37～56℃下干燥过夜。
4. 为了从组织切片上去除石蜡，按照给出的次数和时间，将载玻片依次浸入下面一系列溶剂中。

溶　剂	次数和时间	溶　剂	次数和时间
二甲苯	2×2 min	70%乙醇	1 min
无水乙醇	1 min	用水漂洗	
96%乙醇	1 min		

5. 为了对脱蜡后的组织切片染色，按照给出的时间，将载玻片依次浸入下面一系列溶液中。

溶　液	时　间	溶　液	时　间
Mayer 苏木精	45 s～10 min	用水漂洗 10 min	
或		无水乙醇	1 min
甲基绿	10 s	96%乙醇	1 min
或		70%乙醇	1 min
甲苯胺蓝	10 s	用水漂洗	

6. 为了固定已染色的组织切片，按照给出的次数和时间，将载玻片依次浸入下面一系列溶剂中。

溶　剂	次数和时间	溶　剂	次数和时间
96%乙醇	2×1 min	将载玻片在 37～56℃下干燥过夜	
无水乙醇	2×1 min		

为了得到最佳结果，应该立即进行显微切割。如果有必要，载玻片可以存放数周。通常，标本放置于玻璃载玻片或者薄膜上，并立即用于 LPC 或 LMPC；但是，有一种“液体盖玻片”可以用来改善观察效果和防止标本受到污染，同时仍然可以进行 LMPC 和不受限制的分子分析（见方案 9）。

方案 7

用于 LMPC 的新鲜冰冻组织切片的制备

342

用于 LPC 或 LMPC 的标本应该在液氮中速冻并在－80℃下保存，或者是在－80℃下于 OCT（最佳切片温度复合物）中包埋和保存。如果是后一种情况，在 LPC 之前要从切片中除去 OCT。

如果要进行 RNA 的抽提，切片应在尽可能避免在暴露于 RNase 的条件下进行切片和处理。所有的仪器和染色缸都必须用去除 RNase 污染的试剂，如 RNase*Zap* 处理，并且所有溶液均需用不含 RNase 的成分新鲜配制。更多细节，见第 3 章中 **RNA 操作的特别注意事项**信息栏。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

重要事项：本方案中使用的所有试剂用经过 DEPC 处理的 H_2O 配制（见第 3 章中 **RNA 操作的特别注意事项**信息栏）。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

丙酮（冰冷的）

曙红 Y 的乙醇溶液（HT110-1-16，Sigma Diagnostics，Deisenhofen，Germany）

乙醇（无水的，96％，70％）

据报道，有些批次的乙醇（100％或 200 标准强度）中含有会使微阵列试验中产生很深荧光背景的污染物质，95％的乙醇不存在这个问题。为避免这个问题，可以用 95％的乙醇或已鉴定证明不会产生高荧光背景的无水乙醇。Warner-Graham 公司（Cockeysville，Maryland）出售的无水乙醇（USP 级乙醇，64-17-5）效果很好。

Mayer 苏木精染色液

甲醇（70％）〈!〉

聚-L-赖氨酸（P 1274，Sigma）

RNA*later* 稳定剂（221462，QIAGEN）

可选的，见步骤 3。

RNase*Zap*(9780，Ambion)

细胞和组织

组织标本

专用设备

冰冻切片机

343

玻璃或者膜包被的载玻片（见方案 5）

附加试剂

步骤 1 可能需要方案 5 中所列的试剂和材料。

方法

1. 用冰冻切片机在－20℃下将组织标本切成 10 μm 厚的切片。
2. 将切片直接转移到合适的薄膜载玻片上。
3. 使切片黏附到用膜包被的载玻片上。

   可选的：用 100 μl RNAlater 覆盖在切片上。
4. 为准备用于抽提 RNA 的切片，将载玻片在 70%的乙醇、70% 甲醇或者冰冷的丙酮中浸泡 5 min。

   如果需要的话，可以在此阶段将玻片于－80℃下保存。
5. 为了给切片染色，按照给出的时间，将载玻片依次浸入下面一系列溶液中。

溶　液	时　间	溶　液	时　间
Mayer 苏木精	45 s	70%乙醇	10s
用水漂洗		70%乙醇	10s
用水漂洗(可选的:加入一滴氨水)		96%乙醇	20s
嗜红 Y 的乙醇溶液	20 s	96%乙醇	20s

将载玻片在空气中干燥 30 s 并保存（例如，放在无菌的 50 ml Falcon 试管中）。

为了得到最佳结果，应该立即进行显微切割。如果有必要，载玻片可以放在－80℃下，或者是在干冰上保存过夜。通常，标本放置于玻璃载玻片或者薄膜上，并立即用于 LPC 或 LMPC；但是，有一种“液体盖玻片”可以用来改善观察效果和防止标本受到污染，同时仍然可以进行 LMPC 和不受限制的分子分析（见方案 9）。

344

方案 8

使用 Zincfix 固定组织的 LPC 用于 mRNA 谱分析

对一系列的组织学方法进行了测试以确定它们保留 RNA 的能力，以及与激光显微切割的相容性（Scheidl et al. 2002）。本方案中使用的基于锌的固定剂在保留了 RNA 的同时不会影响到形态学特性。本方案中制备的 RNA 可以用依赖于 RNA 的 DNA 聚合酶转录为高质量的 cDNA 并按照第 3 章中方案 15 和方案 16 介绍的方法进行扩增。RNA 必须用带有噬菌体 T7 启动子（见第 3 章中方案 15 和方案 16）的引物进行扩增（大约 10^6 倍）。为了避免出现问题，应注意防止 RNA 受到 DNA 的污染，特别是含有 T7 启动子的质粒。建议经常清洁实验室的设备，并在自动移液器中使用带滤芯的吸头。

为了获取足够量的 RNA 用于和 cDNA 微阵列进行杂交，在方案的前面部分避免 RNA 的损失非常关键。使用本方案，从组织到分离的总 RNA 总体上的回收估计为大约 50%，损失主要发生在组织的固定和分离步骤。LPC 方法本身和在室温下操作干燥的切片都不会导致 RNA 量的明显减少（Scheidl et al. 2002）。

与本指南中许多其他方案一样，本方案需要一个没有 RNase 污染的工作环境。在适当的时候戴上手套，并在清除污染的台面上工作，也就是说，用诸如 RNase*Zap*（Ambion）试剂处理的工作台面。使用不含 RNase 的酶和试剂。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

重要事项：本方案中使用的所有试剂用经过 DEPC 处理的 H_2O 配制（见第 3 章中 **RNA 操作的特别注意事项**信息栏）。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

乙醇（99.9%，90%，70%）

步骤 11～14 需要 40 ml 每种乙醇溶液。

据报道，有些批次的乙醇（100%或 200 标准强度）中含有会使微阵列试验中产生很深荧光背景的污染物质，95%的乙醇不存在这个问题。为避免这个问题，可以用 95%的乙醇或已鉴定证明不会产生高荧光背景的无水乙醇。Warner-Graham 公司（Cockeysville, Maryland）出售的无水乙醇（USP 级乙醇，64-17-5）效果很好。

OCT（最佳切割温度复合物）

磷酸盐缓冲液（PBS），不含 Ca^{2+} 和 Mg^{2+}

345

RNase*Zap*(9780，Ambion)

溶解在 Zincfix 中的蔗糖（30%，m/V）

二甲苯〈!〉

步骤 11～14 需要 40 ml 二甲苯

Zincfix

5 g $ZnCl_2$

6 g 二水合乙酸锌

0.1 g 醋酸钙

溶解于 1 L 0.1 mol/L Tris-Cl(pH 7.4) 中

重要事项：Zincfix 不稳定，必须新鲜配制。

细胞和组织

组织标本

专用设备

冰冻切片机
用于 OCT 包埋的塑料组织模具
设定为 4℃的摇动平台
用于干燥保存的硅胶
Superfrost Plus 玻璃载玻片（6166 11 55，Fisher Scientific 或者 Menzel Gläser）

方法

切取、固定和包埋

1. 将 PBS 和 Zincfix 在冰上冷却。
2. 在手术后的第一可能时间内，或者在对动物解剖后，将标本转移到冰冷的 PBS 中。
3. 将标本转移入冰冷的 Zincfix 中。

   固定剂：标本的体积比要大于 10。
4. 将标本在 Zincfix 内于 4℃下孵育过夜，并轻轻摇动。
5. 将标本转移入含有 30%(m/V) 蔗糖的 Zincfix 中。在 4℃下孵育标本，直到其沉入容器底部。
6. 将 OCT 倒入组织模具，并将标本按照合适的位置和朝向放入模具中。
7. 将模具放置在干冰上，直到 OCT 冻结。

   将已包埋的标本保存于－80℃，直至切片。

切片

8. 在－20℃下将标本用冰冻切片机切成厚度为 10 μm 的切片。
9. 将切片转移到 Superfrost Plus 玻璃载玻片上。
10. 将切片在室温下干燥 30 min，切片于－80℃下在有硅胶的盒子中保存。

    下面的替代方法可用于得到最佳的组织形态：
    (1) 将 Superfrost Plus 玻片放置在冰冻切片机内数分钟，使其冷却至－20℃。
    (2) 取出每张放在冷的 Superfrost Plus 载玻片上的切片，从玻片的另一侧用一根手指加温使切片黏牢。
    (3) 将切片在冰冻切片机内冻干 30 min。

载玻片的制备

11. 准备 40 ml Zincfix，40 ml 70%乙醇，40 ml 90%乙醇，40 ml 99.5%乙醇和 40 ml二甲苯，装在 50 ml 试管中，并将试管放在冰上。
12. 从－80℃冰箱中取出冰冻的载玻片（来自步骤 10）。
13. 用冰冷的 Zinfix 清洗载玻片 5 min。在本清洗的最后 2 min，用镊子轻轻地摇动载玻片。

14. 将载玻片在下面一系列的脱水剂中浸泡标示的时间：

溶　剂	时　间	溶　剂	时　间
70％乙醇	30 s	99.5％乙醇	30 s
95.5％乙醇	30 s	二甲苯	1 min

15. 将载玻片于室温下在通风橱内干燥。

为了得到最佳结果，应该立即进行显微切割。如果有必要，载玻片可以放置 18～20 h (Scheidl et al. 2002)。通常，标本放置于玻璃载玻片或者薄膜上，并立即用于 LPC 或 LMPC；但是，有一种“液体盖玻片”可以用来改善观察效果和防止标本受到污染，同时仍然可以进行 LMPC 和不受限制的分子分析（见方案 9）。

方案 9

使用液体盖玻片保护组织样品

为了进行显微分析，组织学切片通常包埋在“支撑”介质中并盖上盖玻片。但是，固定介质和盖玻片的存在影响了使用激光压力弹射对样品的回收。非常遗憾的是，省略介质和盖玻片会影响光的折射并降低精确鉴定靶细胞所需要的图像的质量。有时可以通过常规的组织学或者荧光染色切片来缓解这一问题。当这些办法都不合适时，靶细胞必须依靠其形态学特征进行鉴别。在这些情况下，建议使用能够改善光折射的支撑介质。要在短时间内改善观察效果，可以将 70％乙醇或者 1％甘油的 PBS 溶液加到载玻片上的切片中。这些标本在激光切割之前要进行干燥。

PALM 评价了一种改进的“液体盖玻片”试剂，可以在切片上聚合，显示的折射率与组织相似，并且对后续的分子分析没有影响。这种“液体盖玻片”同时也起着保护样品不受污染的作用。下面的方案，由 Tone Bjornsen 和 Christer Bush（University Hospital，Department of Pathology，Tromsoe，Norway）提供，该方案基于一种无毒害的溶液，并且容易喷洒到载玻片上。

“PALM 液体盖玻片”也可以用于 LMPC，并且不会影响到分子分析。

材料

重要事项：本方案中使用的所有试剂用经过 DEPC 处理的 H_2O 配制（见第 3 章中 **RNA 操作的特别注意事项**信息栏）。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

异丙醇

液体盖玻片溶液（PLCS，PALM）

松节油可以用来清洁或者去除不需要的残留 PLCS。

细胞和组织

裱贴在载玻片上的组织切片，如方案 6、方案 7 或者方案 8 中的介绍。

专用设备

喷雾瓶

348 ## 方法

重要事项：在化学通风橱中进行全部的操作步骤以减少污染。

1. 用异丙醇按照 1∶5 或者 1∶10 的比例稀释液体盖玻片溶液。
2. 用喷雾瓶向裱在膜上的切片表面加上一薄层稀释的溶液。
3. 将载玻片放在化学通风橱中，让异丙醇从切片上挥发。
4. 进行显微切割和激光压力弹射。

   这些切片可以存放于室温下直至后面的处理。

349 # 方案 10

激光显微切割和弹射后活细胞的分离

激光微束显微切割（laser microbeam microdissection，LMM）和激光压力弹射（LPC）是用于从固定或冰冻组织中提取细胞的方法（Meier-Ruge et al. 1976；Schindler et al. 1985；Kubo et al. 1995；Emmert-Buck et al. 1996；Bonner et al. 1997；Schiitze and Lahr 1998；Schütze et al. 1998；Lahr 2000）。在此之前，基于激光的活细胞俘获通常会造成细胞的破坏（Schindler et al. 1985）；经过改进的 LMPC 激光技术和新近发展起来的细胞培养方法使得单个活细胞或者成团细胞的显微切割和“弹射”成为可能（Mayer et al. 2002）。这个方案用于那些细胞培养物，其中选定的细胞被显微切割、弹出并回收用于进一步的增殖。俘获和弹射的细胞的活力不会受到本方法的影响。

为了对活细胞进行 LMPC，将细胞培养在一个组合的“圆形开口和密闭的”（round open closed，ROC）小室之中，或者是在 PALM DuplixDishes 中。使用 20× 物镜勾勒出单个细胞或者是成团的细胞。来自勾勒出的区域中所不希望的细胞可直接用激光照射来清除。进行 LPC 时，单次激光照射作用于勾勒出的薄膜-细胞堆叠层的边缘部分。即使是大的细胞-薄膜堆叠层（如 385 μm×248 μm）也可以被弹射（图 5-11）。单个活细胞的显微切割在数秒内完成，而且成团细胞的弹射不到 2 min。被分离和弹射的细胞悬浮在合适体积的补料和条件培养液中。如果分离细胞的数目不足十个，最好将它们放在所谓的“悬滴”中并保持密切接触。细胞（仍然是球形的）在悬滴的底部聚集，并开始形成细胞-细胞之间的接触（图 5-12C）。

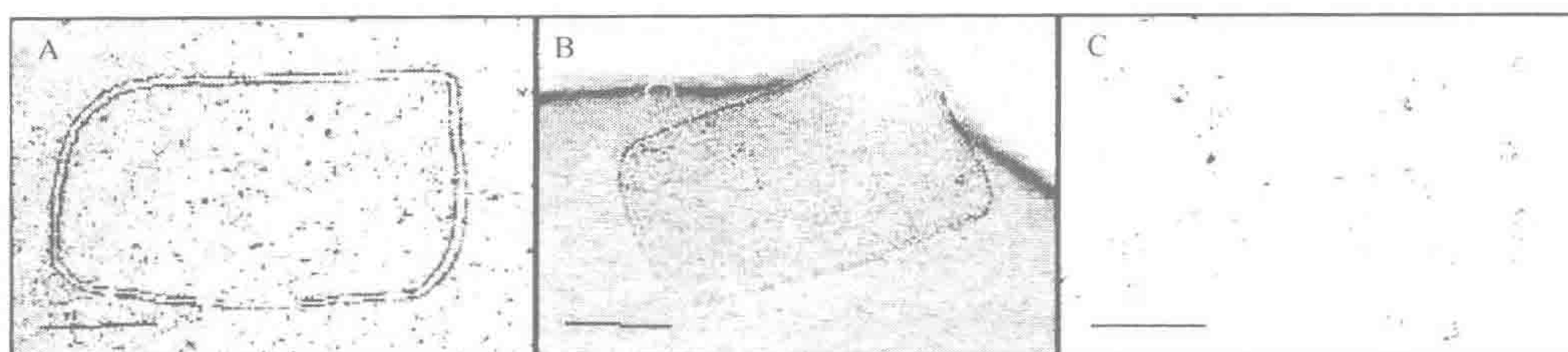


图 5-11 使用 LMPC 俘获 EJ28 细胞的显微图解

图像显示了显微切割之后的细胞（A），弹射的带有细胞的薄膜（B），和培养 1 天之后的细胞（C）。

比例尺，100 μm。物镜，20×

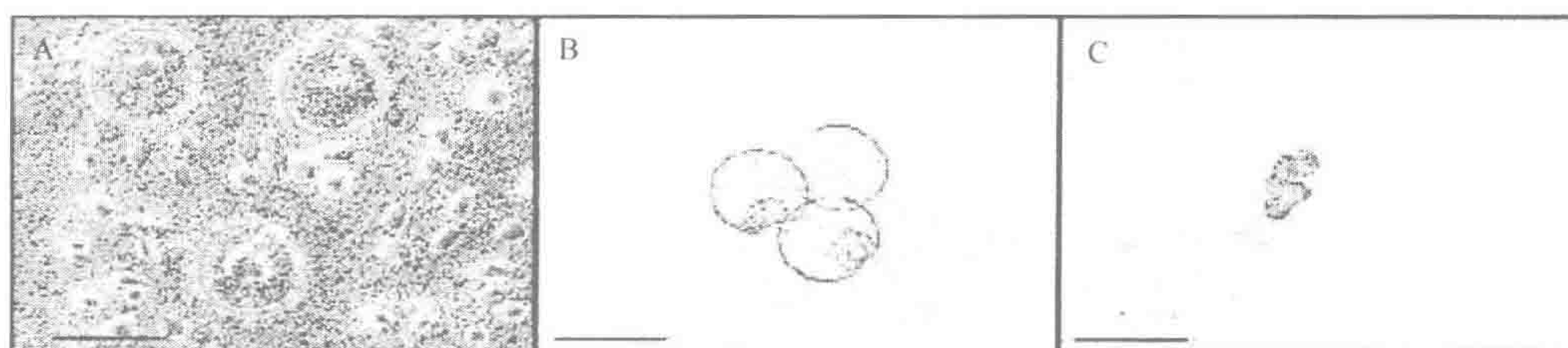


图 5-12 合并的单个 EJ28 细胞的图像

A. 激光切割后的细胞；B. 三个被弹射的、带有细胞的膜区域；C. 在“悬滴”中聚集的细胞。

比例尺：（A，B）100 μm；（C）50 μm。物镜：（A，B）20×；（C）40×

350

第二天，聚集的细胞平铺于标准培养皿底部的一小滴补料培养液中。由于小滴外形是球形的，聚集的细胞被迫迁移至没有张力的液滴中心，并且经过 1 天的培养，细胞恢复了它们典型的形态学特征（图 5-11 C）。根据平铺细胞的密度和它们在显微切割之前的生理状态，在 3 天（30 个细胞）至最多 3 周（10 个细胞）的时间内开始增殖。

LMPC 这种使用精确的激光照射以消除不希望的细胞和能够回收活细胞的能力，能够用来从不均一的细胞培养物（如在转染实验之后）中建立均一的细胞种群。理论上，这个方法可以进一步改进用于从活检得到的组织中分离活细胞。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

乙醇（无水的）

Hank 平衡盐溶液（HBSS）（H 6136，Sigma-Aldrich）

胰蛋白酶-EDTA 溶液（Sigma-Aldrich）

细胞培养液

含有 1× 抗生素和抗霉素溶液的合适的细胞培养液（Sigma-Aldrich）

用于所希望细胞系的条件培养液（Patterson and Chun 1977；Garcia-Moreno et al. 2002）

补料的培养液（4 份培养液，1 份条件培养液）

专用设备

用于细胞培养的培养皿（盘和盖子）

用于细胞培养的 CO_2 培养箱（Kendro Laboratory Products）

倒置显微镜（Axiovert 135，Zeiss）

激光显微镜，Robot-MicroBeam（PALM Microlaser Technologies AG，Bernried，Germany）

微量离心管（PALM Microlaser Technologies）

聚萘二甲酸乙二醇酯薄膜（PALM Microlaser Technologies）

ROC（圆形开口和密闭的）Chamber，（PeCon and LaCon，Erbach-Bach，Germany）或者 DuplixDishes(PALM Microlaser Technologies)

紫外灯，专用于在安全罩中使用（Kendro Laboratory Products）〈!〉

方法

重要事项：为了减少污染的机会，在操作细胞培养物时需戴上手套。不要让细胞培养物在培养箱外停留超过需要的时间。

351

1. 按照制造商的说明装配 ROC 小室，但要另外加入一张 PEN 薄膜。
2. 向小室底部的玻片上加入一滴无水乙醇，以便于膜的放置。
3. 在 ROC 小室底部的玻片上盖上薄膜。
4. 将开口的带有薄膜的小室用紫外灯照射 20 min，使薄膜特性从疏水转变为亲水。
5. 完成 ROC 小室的装配，将其在 121℃下高压灭菌 20 min（高压蒸汽灭菌）。

细胞培养

6. 将培养细胞以所希望的密度接种于位于 ROC 小室底部覆盖的膜上的合适培养液中，理想情况下应该形成半汇合的单层。细胞在气体培养箱中于 37℃下培养1～2 天。此时细胞即可用于显微切割。

激光显微切割和弹射

7. 在进行激光显微切割之前，从 ROC 小室中彻底清除培养液。
8. 根据 PALM 提供的信息材料中的介绍，显微切割位于膜上的所希望的细胞。

   激光切割参数取决于特定的激光-显微镜系统，并且必须优化。
9. 小心地向选定的细胞上面加上 10 μl 的 HBSS 以便于 LPC，注意不要冲走显微切割的标本。
10. 吸取 5 μl 的 HBSS 并加到微量离心管盖子的中央，然后在不使用操作器的情形下将盖子直接放在 ROC 小室中选定细胞的上方。

或者，也可以用 PALM Petriperm 夹持器联合 CapMover 来俘获细胞。

11. 使用单次激光照射，作用于勾勒出的薄膜-细胞堆叠层的边缘来弹射细胞薄膜堆叠层。

    能量的设定要足够地高，以便将带有膜的显微切割的细胞弹射到盖子中，甚至大的细胞-薄膜堆叠层（如 385 μm×248 μm）也能够被弹射。

12. PLC 结束后，从显微镜的载物台上移开 ROC 小室并观察盖子中弹射的细胞。

    如果要进行多次实验，向细胞中加入培养液使 ROC 小室中的细胞从干燥状态得到复原，然后按照从步骤 7 开始的介绍重复显微切割。数次显微切割以后，液体/培养液进入 ROC 小室玻璃基底和薄膜之间的微小空隙，这使得进一步的显微切割和弹射变得越来越困难，并且最终不能继续进行。

弹射细胞的回收

352

13. 向盖子内弹射的细胞加入 25 μl 的 HBSS。
14. 将倒转的试管放到盖子上，使得细胞被密封。将细胞在室温下保存最多 30 min，以使它们与膜分离。
15. 将试管于室温下，在微量离心机中以 8000 *g* 离心 1 min，弃去上清以回收细胞。
16. 将细胞重悬于 20 μl 胰蛋白酶-EDTA 溶液，在室温下温育 10 min。
17. 将细胞于室温下，在微量离心机中以 8000 *g* 离心 1 min。
18. 离心后，将胰蛋白酶处理的细胞重悬于 15 μl(＜10 个细胞) 或者 20 μl 补料的培养液（4 份培养液，1 份条件培养液）。

    如果试管中弹射的细胞少于 10 个，继续步骤 19；如果悬液中有 10 个或者更多的细胞，前进到步骤 23。

细胞的培养

19. 如果悬液中回收的细胞少于 10 个，将含有细胞的 15 μl 液滴移到倒置的培养皿盖内侧形成“悬滴”。
20. 翻转盖子，并将其放入培养皿的底部。这使得“悬滴”中的单个细胞之间产生密切接触并开始聚集。
21. 将“悬滴”中的细胞在气体培养箱中于 37℃下培养过夜。
22. 培养后，再次翻转盖子，并向液滴中加入 5 μl 的补料培养液（共 20 μl)。
23. 用灭菌的微量移液管将 2 μl 含有细胞的悬液加入培养皿的底部，可以是单独的或者与其他液滴之间有足够的空间。为了避免单个液滴之间细胞发生交叉污染的可能性，用硅胶或者克隆环圈出每一群细胞。
24. 将细胞在 37℃下培养过夜
25. 培养后，使用永久性记号笔在培养皿底部标记出液滴，以便于对细胞的追踪。
26. 去除培养液，根据培养皿直径的大小，向培养皿中加入 4～10 ml 新鲜的补料培养液。

    在随后的数日或者数周中可以观察到细胞的生长。建议经常更换细胞培养液。

Banks R.E., Dunn M.J., Forbes M.A., Stanley A., Pappin D., Naven T., Gough M., Harnden P., and Selby P.J. 1999. The potential use of laser capture microdissection to selectively obtain distinct populations of cells for proteomic analysis—Preliminary findings. *Electrophoresis* **20:** 689–700.

Becker I., Becker K.F., Rohrl M.H., and Hofler H. 1997. Laser-assisted preparation of single cells from stained histological slides for gene analysis. *Histochem. Cell Biol.* **108:** 447–451.

Becker I., Becker K.F., Rohrl M.H., Minkus G., Schütze K., and Hofler H. 1996. Single-cell mutation analysis of tumors from stained histologic slides. *Lab. Invest.* **75:** 801–807.

Bernsen M.R., Dijkman H.B., de Vries E., Figdor C.G., Ruiter D.J., Adema G.J., and van Muijen G.N. 1998. Identification of multiple mRNA and DNA sequences from small tissue samples isolated by laser-assisted microdissection. *Lab. Invest.* **78:** 1267–1273.

Bertheau P., Plassa L.F., Lerebours F., de Roquancourt A., Turpin E., Lidereau R., de Thé H., and Janin A. 2001. Allelic loss detection in inflammatory breast cancer: Improvement with laser microdissection. *Lab Invest.* **81:** 1397–1402.

Böhm M. and Wieland I. 1997. Analysis of tumour-specific alterations in native specimens by PCR: How to procure the tumour cells. *Int. J Oncol.* **10:** 11–139.

Böhm M., Wieland I., Schütze K., and Rübben H. 1997. MicroBeam MOMeNT—Non contact laser microdissection of membrane-mounted native tissue. *Am. J. Pathol.* **151:** 63–67.

Bonner R.F., Emmert-Buck M., Cole K., Pohida T., Chuaqui R., Goldstein S., and Liotta L.A. 1997. Laser capture microdissection: Molecular analysis of tissue. *Science* **278:** 1481–1483.

Buerger H., Gebhardt F., Schmidt H., Beckmann A., Hutmacher K., Simon R., Lelle R., Boecker W., and Burkhard B. 2000a. Length and loss of heterozygosity of an intron 1 polymorphic sequence of *egfr* is related to cytogenetic alterations and epithelial growth factor receptor expression. *Cancer Res.* **60:** 854–857.

Buerger H., Simon R., Schafer K.L., Diallo R., Littmann R., Poremba C., van Diest P.J., Dockhorm-Dworniczak B., and Boecker W. 2000b. Genetic relation of lobular carcinoma in situ, ductal carcinoma in situ, and associated invasive carcinoma of the breast. *Mol. Pathol.* **53:** 118–121.

Burgemeister R., Schütze K., Minderer S., and Gloning K.-P. 1999. Single fetal cells separated by the Laser MicroBeam technique. *BIOforum Int.* **3:** 119–121.

Burton M.P., Schneider B.G., Brown R., Escamilla-Ponce N., and Gulley M.L. 1998. Comparison of histologic stains for use in PCR analysis of microdissected, paraffin-embedded tissues. *BioTechniques* **24:** 86–92.

Dietmaier W., Gänsbauer S., Beyser K., Renke B., Hartmann A., Rümmele P., Jauch K.-W., Hofstädter F., and Rüschoff J. 2000. Microsatellite instability in tumor and nonneoplastic colorectal cells from hereditary non-polyposis colorectal cancer and sporadic high microsatellite-instable tumor patients. *Pathobiology* **68:** 227–231.

Emmert-Buck M.R., Bonner R.F., Smith P.D., Chuaqui R.F., Zhuang Z., Goldstein S.R., Weiss R.A., and Liotta L.A. 1996. Laser capture microdissection. *Science* **274:** 998–1001.

Emmert-Buck M.R., Lubensky I.A., Chuaqui R.F., Debelenko L.V., Vocke C.D., Merino M.J., Duray P.H., Linehan W.M., Liotta L.A., and Zhuang Z. 1998. Applications of tissue microdissection in molecular pathology: Principles and guidelines. *Methods Mol. Med.* **14:** 269–298.

Emmert-Buck M.R., Gillespie J.W., Paweletz C.P., Ornstein D.K., Basrur V., Appella E., Wang Q.H., Huang J., Hu N., Taylor P., and Petricoin III E.F. 2000. An approach to proteomic analysis of human tumors. *Mol. Carcinog.* **27:** 158–165.

Fend F., Emmert-Buck M.R., Chuaqui R., Cole K., Lee J., Liotta L.A., and Raffeld M. 1999a. Immuno-LCM: Laser capture microdissection of immunostained frozen sections for mRNA analysis. *Am. J. Pathol.* **154:** 61–66.

Fend F., Quintanilla-Martinez L., Kumar S., Beaty M.W., Blum L., Sorbara L., Jaffe E.S., and Raffeld M. 1999b. Composite low grade B-cell lymphomas with two immunophenotypically distinct cell populations are true biclonal lymphomas. A molecular analysis using laser capture microdissection. *Am. J. Pathol.* **154:** 1857–1866.

Fink L., Seeger W., Ermert L., Hänze J., Stahl U., Grimminger F., Kummer W., and Bohle R.M. 1998. Real-time quantitative RT-PCR after laser-assisted cell picking. *Nat. Med.* **11:** 1329–1333.

Fink L., Kohlhoff S., Stein M.M., Hanze J., Weissmann N., Rose F., Akkayagil E., Manz D., Grimminger F., Seeger W., and Bohle R.M. 2002. cDNA array hybridization after laser-assisted microdissection from nonneoplastic tissue. *Am. J. Pathol.* **160:** 81–90.

Garcia-Moreno C., Mendez-Davila C., de La Piedra C., Castro-Errecaborde N.A., and Traba M.L. 2002.

Human prostatic carcinoma cells produce an increase in the synthesis of interleukin-6 by human osteoblasts. *Prostate* **50:** 241–246.

Gillespie J.W., Ahram M., Best C.J., Swalwell J.I., Krizman D.B., Petricoin E.F., Liotta L.A., and Emmert-Buck M.R. 2001. The role of tissue microdissection in cancer research. *Cancer J.* **7:** 32–39.

Gillespie J.W., Best C.J.M., Bischel V.E., Cole K.A., Greenhut S.F., Hewitt S.M., Ahram M., Gathright Y.B., Merino M.J., Strausberg R.L, Epstein J.I., Hamilton S.R., Gannot G., Baibakova G.V., Calvert V.S., Flaig M.J., Chuaqui R.F., Herring J.C., Pfeifer J., Petricoin, E.F., Linehan W.M., Duray P.H., Bova G.S., and Emmert-Buck M.R. 2002. Evaluation of non-formalin tissue fixation for molecular profiling studies. *Am. J. Pathol.* **160:** 449–457.

Gjerdrum L.M., Lielpetere I., Rasmussen L.M., Bendix K., and Hamilton-Dutoit S. 2001. Laser-assisted microdissection of membrane-mounted paraffin sections for polymerase chain reaction analysis: Identification of cell populations using immunohistochemistry and in situ hybridization. *J. Mol. Diagn.* **3:** 105–110.

Glöckner S., Lehmann U., Wilke N., Kleeberger W., Länger F., and Kreipe H. 2000. Detection of gene amplification in intraductal and infiltrating breast cancer by laser-assisted microdissection and quantitative real-time PCR. *Pathobiology* **68:** 173–179.

Glöckner S., Lehmann U., Wilke N., Kleeberger W., Länger F., and Kreipe H. 2001. Amplification of growth regulatory genes in intraductal breast cancer is associated with higher nuclear grade but not with the progression to invasiveness. *Lab Invest.* **81:** 565–571.

Going J.J. and Lamb R.F. 1996. Practical histological microdissection for PCR analysis. *J. Pathol.* **179:** 121–124.

Goldsworthy S.M., Stockton P.S., Trempus C.S., Foley J.F., and Maronpot R.R. 1999. Effects of fixation on RNA extraction and amplification from laser capture microdissected tissue. *Mol. Carcinog.* **25:** 86–91.

Heel K. and Dawkins H. 2001. Laser microdissection and optical tweezers in research. *Today's Life Sci.* March/April: 42–48.

Hiller T., Snell L., and Watson P.H. 1996. Microdissection RT-PCR analysis of gene expression in pathologically defined frozen tissue sections. *BioTechniques* **21:** 38–42.

Hoffmann M., Olson K., Cavender A., Pasqualini R., Gaikwad J., and D'Souza R.N. 2001. Gene expression in a pure population of odontoblasts isolated by laser-capture microdissection. *J. Dent. Res.* **80:** 1963–1967.

Imamichi Y., Lahr G., and Wedlich D. 2001. Laser-mediated microdissection of paraffin sections from *Xenopus* embryos allows detection of tissue-specific expressed mRNAs. *Dev. Genes Evol.* **211:** 361–366.

Kaserer K., Schmaus J., Bethge U., Migschitz B., Fasching S., Walch A., Herbst F., Teleky B., and Wrba F. 2000. Staining patterns of p53 immunohistochemistry and their biological significance in colorectal cancer. *J. Pathol.* **190:** 450–456.

Kleeberger W., Rothämel T., Glöckner S., Lehmann U., and Kreipe H. 2000. Laser-assisted microdissection and short tandem repeat PCR for the investigation of graft chimerism after solid organ transplantation. *Pathobiology* **68:** 196–201.

Kubo Y., Klimek F., Kikuchi Y., Bannasch P., and Hino O. 1995. Early detection of Knudson's two-hits in preneoplastic renal cells of the Eker rat model by the laser microdissection procedure. *Cancer Res.* **55:** 989–990.

Lahr G. 2000. RT-PCR from archival single cells is a suitable method to analyse specific gene expression. *Lab Invest.* **80:** 1477–1479.

Lahr G., Stich M., Schutze K., Blumel P., Posl H., and Nathrath W.B. 2000. Diagnosis of papillary thyroid carcinoma is facilitated by using an RT-PCR approach on laser-microdissected archival material to detect RET oncogene activation. *Pathobiology* **68:** 218–226.

Lehmann U., Glöckner S., Kleeberger W., Feist H., von Wasielewski R., and Kreipe H. 2000. Detection of gene amplification in archival breast cancer specimens by laser-assisted microdissection and quantitative real-time polymerase chain reaction. *Am. J. Pathol.* **156:** 1855–1864.

Ling G., Persson A., Berne B., Uhlén M., Lundeberg J., and Pónten F. 2001. Persistent p53 mutations in single cells from normal human skin. *Am. J. Pathol.* **159:** 1247–1253.

Luo L., Salunga R.C., Guo H., Bittner A., Joy K.C., Galindo J.E., Xiao H., Rogers K.E., Wan J.S., Jackson M.R., and Erlander M.G. 1999. Gene expression profiles of laser-captured adjacent neuronal subtypes. *Nat. Med.* **5:** 117–122.

Mayer A., Stich M., Brocksch D., Schütze K., and Lahr G. 2002. Going *in vivo* with laser-microdissection. *Methods Enzymol.* (in press).

Meier-Ruge W., Bielser W., Remy E., Hillenkamp F., Nitsche R., and Unsöld R. 1976. The laser in the Lowry technique for microdissection of freeze-dried tissue slices. *Histochem. J.* **8:** 387–401.

355

Nagasawa Y., Takenaka M., Matsuoka Y., Imai E., and Hori M. 2000. Quantitation of mRNA expression in glomeruli using laser-manipulated microdissection and laser pressure catapulting. *Kidney Int.* **57:** 717–723.

Noack F., Helmecke D., Rosenberg R., Thorban S., Nekarda H., Fink U., Lewald J., Schütze K., Stich M., Harbeck N., Magdolen V., Graeff H., and Schmitt M. 1999. CD87-positive tumor cells in bone marrow aspirates identified by confocal laser scan fluorescence microscopy. *Int. J. Oncol.* **15:** 617–623.

Ohyama H., Zhang X., Kohno Y., Alevizos I., Posner M., Wong D.T., and Todd R. 2000. Laser capture microdissection-generated target sample for high-density oligonucleotide array hybridization. *BioTechniques* **29:** 530–536.

Ornstein D.K., Gillespie J.W., Paweletz C.P., Duray P.H., Herring J., Vocke C.D., Topalian S.L., Bostwick D.G., Linehan W.M., Petricoin III E.F., and Emmert-Buck M.R. 2000. Proteomic analysis of laser capture microdissected human prostate cancer and in vitro prostate cell lines. *Electrophoresis* **21:** 2235–2242.

Patterson P.H. and Chun L.L. 1977. The induction of acetylcholine synthesis in primary cultures of dissociated rat sympathetic neurons. II. *Dev. Biol.* **60:** 473–481.

Persson A.E., Ling G., Williams C., Bäckvall H., Pontén J., Pontén F., and Lundeberg J. 2000. Analysis of p53 mutations in single cells obtained from histological tissue sections. *Anal. Biochem.* **287:** 25–31.

Peterson L.A., Brown M.R., Carlisle A.J., Kohn E.C., Liotta L.A., Emmert-Buck M.R., and Krizman D.B. 1998. An improved method for construction of directionally cloned cDNA libraries from microdissected cells. *Cancer Res.* **58:** 5326–5328.

Ponten F., Williams C., Ling G., Ahmadian A., Nister M., Lundeberg J., Ponten J., and Uhlen M. 1997. Genomic analysis of single cells from human basal cell cancer using laser-assisted capture microscopy. *Mutat. Res.* **382:** 45–55.

Scheidl S.J., Nilsson S., Kalèn M., Hellström M., Takemoto M., Hakansson J., and Lindahl P. 2002. mRNA expression profiling of laser microbeam microdissected cells from slender embryonic structures. *Am. J. Pathol.* **160:** 801–813.

Schindler M. 1998. Select, microdissect, and eject. *Nat. Biotechnol.* **16:** 719–720.

Schindler M., Allen M.L., Olinger M.R., and Holland J.F. 1985. Automated analysis and survival selection of anchorage-dependent cells under normal growth conditions. *Cytometry* **6:** 368–374.

Schütze K. and Lahr G. 1998. Identification of expressed genes by laser-mediated manipulation of single cells. *Nat. Biotechnol.* **16:** 737–742.

Schütze K., Pösl H., and Lahr G. 1998. Laser micromanipulation systems as universal tools in molecular biology and medicine. *Mol. Cell. Biol.* **44:** 735–746.

Shibata D., Hawes D., Li Z.H., Hernandez A.M., Spruck C.H., and Nichols P.W. 1992. Specific genetic analysis of microscopic tissue after selective ultraviolet radiation fractionation and the polymerase chain reaction. *Am. J. Pathol.* **141:** 539–543.

Simone N.L., Bonner R.F., Gillespie J.W., Emmert-Buck M.R., and Liotta L.A. 1998. Laser-capture microdissection: Opening the microscopic frontier to molecular analysis. *Trends Genet.* **14:** 272–276.

Specht K., Richter T., Müller U., Walch A., Werner M., and Höfler H. 2001. Quantitative gene expression analysis in microdissected archival formalin-fixed and paraffin-embedded tumor tissue. *Am. J. Pathol.* **158:** 419–429.

Srinivasan R. 1986. Ablation of polymers and biological tissue by ultraviolet lasers. *Science* **234:** 559–565.

Stoehr R., Hartmann A., Hindlmeyer E., Mürle K., Wieland W., and Knuechel R. 2000. Oligoclonality of early lesions of the urothelium as determined by microdissection-supported genetic analysis. *Pathobiology* **68:** 165–172.

Suarez-Quian C.A., Goldstein S.R., Pohida T., Smith P.D., Peterson J.I., Wellner E., Ghany M., and Bonner R.F. 1999. Laser capture microdissection of single cells from complex tissues. *BioTechniques* **26:** 328–335.

Vona G., Sabile A., Louha M., Sitruk V., Romana S., Schütze K., Capron F., Franco D., Pazzagli M., Vekemans M., Lacour B., Bréchot C., and Paterlini-Bréchot P. 2000. Isolation by size of epithelial tumor cells. A new method for the immunomorphological and molecular characterization of circulating tumor cells. *Am. J. Pathol.* **156:** 57–63.

Walch A., Specht K., Smida J., Aubele M., Zitzelsberger H., Höfler H., and Werner M. 2001. Tissue microdissection techniques in quantitative genome and gene expression analyses. *Histochem. Cell Biol.* **115:** 269–276.

Westphal G., Burgemeister R., Friedemann G., Wellmann A., Wernert N., Wollscheid V., Becker B., Vogt T., Knüchel R., Stolz W., and Schütze K. 2002. Non-contact laser catapulting—A basic procedure for functional genomics and proteomics. *Methods Enzymol.* (in press).

因特网资源

http://cgap-mf.nih.gov/index.html Molecular Profiling Initiative
http://www.arctur.com Arcturus Systems for Microgenomics homepage
http://www.arctur.com/products/picopure_rna_isolation.htm Arcturus PicoPureM RNA isolation Kit
http://www.PALM-Microlaser.com P.A.L.M. Mirolaser Technologies homepage
http://www.qiagen.com/literature/brochures/pcr/pdf/pcrcha28.pdf DiLella A.G. 2001. Department of Pharmacology, Merck Research Laboratories. Analysis of RNA from laser capture microdissected cells.
http://www.stratagene.com/cellbio/purification/absolutely_rna_micro.htm Stratagene Absolutely RNA Microprep Kit

357

第6章　基因组微阵列分析

导言

PART Ⅰ　微阵列检测 DNA 拷贝数

导言

方案

1. cDNA 微阵列比较基因组杂交
2. 制备 BAC DNA 用于通过连接介导 PCR 进行 CGH
3. BAC 基因组微阵列 CGH
4. 基于微阵列的 DNA 拷贝数代表性分析：靶 DNA 制备
5. 基于微阵列的 DNA 拷贝数代表性分析：玻片制备及杂交

PART Ⅱ　突变检测及 SNP 基因分型

导言

方案

6. SBE-标签阵列用于 SNP 基因分型
   - 替代方案：为 SBE 进行多重扩增反应
7. Affymetrix 标签阵列用于 SBE SNP 基因分型
8. 高密度微阵列用于序列变异检测
9. SNP 阵列用于杂合性缺失分析

PART Ⅲ　基于微阵列的 DNA-蛋白质相互作用检测：微阵列上的染色质免疫沉淀

导言

方案

358

10. 应用阵列对酵母基因组上 DNA-蛋白质相互作用作图
11. 酵母基因组的 DNA-蛋白质相互作用作图：DNA 扩增及标记

PART Ⅳ　SNP 序列：作图数据

- 表 6-7 多重引物集合：SNP 作图数据
- 表 6-8 Affymetrix HuSNP 芯片中的 SNP 在图谱上的位置

参考文献

因特网资源

导言

目前，许多生物大量的序列及多态性数据已经获得，这为基因组结构及功能的研究提供了异常灵敏的工具。在这本指南前面的章节里，已经讨论了使用微阵列研究特定细胞类型及组织样品的基因表达谱。本章中汇集的方案提供了基于微阵列的，用于三种不同类型基因组分析的策略：基因组拷贝数的确定、多态性检测及作图、基因组 DNA-蛋白质相互作用作图。

本章的 PART Ⅰ描述了利用微阵列估测 DNA 拷贝数的三种方法。在前两种方法中，利用代表了基因组 DNA 全部复杂性的 cDNA 克隆（方案 1）或细菌人工染色体（bacterial artificial chromosome，BAC）克隆（方案 2 和方案 3）进行比较基因组杂交（comparative genomic hybridization，CGH）。在第三种方法中，使用了一种称为代表性差异分析（representational difference analysis，RAD）的方法，应用基因组代表使得总基因组 DNA 的序列复杂性显著降低。

PART Ⅱ提供了关于突变和（或）基因组序列变异的检测及作图方案。代表着特定已知序列的寡核苷酸探针进行有序排布的可重复阵列，非常适合哺乳动物基因组的作图及分析。最近，在人类及其他基因组中大量单核苷酸探多态性（single-nucleotide polymorphisms，SNP）的发现和鉴定，提供了高密度覆盖基因组的，广泛的物理标记图谱。这些特征为检测特定基因组区域中新的自发突变或鉴定遗传性 SNP 的不同方法奠定了基础。PART Ⅱ中的方案描述了使用单碱基延伸（single-base extension，SBE）——标签阵列（方案 6）或通用标签阵列（方案 7）检测 SNP、重测序（resequencing）检测序列变异（方案 8）和使用 SNP 阵列分析杂合性缺失（LOH）（方案 9）。

最后，第三部分描述了使得 DNA-蛋白质相互作用作图更为便利的微阵列杂交方法（方案 10 和方案 11）。这些方法建立在使用基本上代表了每个酵母基因间隔（启动子）序列的微阵列的基础上。从特异性免疫沉淀染色体中获得的 DNA 被扩增，标记并与阵列杂交。尽管这一方法是为酵母基因组而建立的，但是能够在改造后适用于任何序列完全确定的基因组，或者用于探询基因组特定的区域。

关于这些分析方法中每种方法的更多信息，请见本章各部分的导言。

PART Ⅰ 微阵列检测 DNA 拷贝数

359

导言

比较基因组杂交（comparative genomic hybridization，CGH）可以在单次测定中鉴定和绘制出全基因组上 DNA 拷贝数变异的位点（Kallioniemi et al. 1992，1993；du Manoir et al. 1993）。在最初建立的 CGH 中，来自两个（或更多）细胞种群的总基因组 DNA 用不同的荧光染料标记，并与正常个体的中期铺展染色体进行杂交。可从新鲜、冷冻或固定组织样品中获得基因组 DNA。加入高浓度的未标记重复 DNA，通常是 C_0t-1 DNA，以封闭重复序列的杂交。来自不同细胞种群的序列与其染色体上互补位置

结合的比率，与它们在杂交混合物中的相对初始浓度成比例。因此，定量计算在基因组图谱中染色体上的杂交比率，提供了不同序列在被杂交基因组中相对拷贝数的信息。如果一个细胞种群有“正常”的二倍体基因组，通常称为“参考”，那么从不同位点发射的信号的比率就提供了在“实验”基因组中拷贝数变异的图谱。

DNA 拷贝数变异对生殖及机体健康问题的重要作用促进了 CGH 的发展。例如，特定基因组位点上组成性缺失和加倍常常与胎儿发育异常有关，导致唐氏综合征等情况。在癌症中，恶性发展伴随着体细胞基因组中出现多重改变。这些改变中的一部分包括 DNA 拷贝数的局部丢失或增益，因而能够被 CGH 检测出来。因为丢失会降低基因的表达，它们促成，如抑癌基因的失活。反之，拷贝数的增益能够升高促使细胞生长的基因的表达水平。观察到的一些拷贝数增益来源于含有已知癌基因的基因组焦点区域的高水平扩增；其他的可能预示着新的癌基因的位置。

自应用以来，CGH 提供了有关在癌细胞系和肿瘤样品中 DNA 扩增和缺失区域的丰富信息（见 http：//www. progenetix. net/的出版物）。CGH 分析同样直接为鉴定在肿瘤发生或发展过程中起着重要作用的基因作出了贡献（见 Visakorpi et al. 1995a，b；Monni et al. 1996；Hem min ki et al. 1997；Collins et al. 1998）。然而，使用中期染色体作为作图骨架导致的低的细胞遗传学分辨率限制了作图的精确性和对微小异常的检测。此外，尽管有限的基因组区域的高水平扩增易于检测，但是两种不同样品发射信号的比率不一定就反映了拷贝数增加的倍数。

用微阵列进行 CGH（Solinas-Toldo et al. 1997；Pinkel et al. 1998；Pollack et al. 1999）能够克服很多传统的细胞遗传学 CGH 的限制。

360

使用已作图基因组克隆构成的阵列允许测量的分辨率由阵列上克隆间的间距和它们的大小来确定。然而，精确测量拷贝数这种性能上的严格要求使得阵列 CGH 成为一种要求非常高的微阵列技术。例如，一个人希望能够在由一种以上类型细胞组成的样本，如肿瘤和正常细胞的混合物中检测生物学相关的单拷贝 DNA 的增益和丢失。这只有当 CGH 足够灵敏，能够可靠检测小于 2 倍的比率变化时才能实现。在哺乳动物系统中，这个精度要能够在核酸制备物，基因组 DNA 中实现，其复杂性至少是相应的全体 mRNA的 10 倍。

第一部分满足了检测 DNA 拷贝数变化的需要。在前两种方法中，总基因组 DNA 与包含下列材料的任意一种阵列杂交：

- 聚合酶链反应（PCR）扩增的 cDNA（方案 1）（Pollsck et al. 1999）。同一个 cDNA 微阵列可以用于进行表达分析和基因组 DNA 分析，这在拥有高质量 cDNA阵列时是一个明显的好处。对许多在染色体上相邻的阵列中的 cDNA 的荧光比率进行平均，能够容许对单拷贝增益和丢失进行较高质量的检测。
- 在高容量载体，如 P1 和 BAC 中构建的基因组文库（方案 2 和 3）。使用包含非常大的基因组插入片段的阵列（Salinas-Toldo et al. 1997；Pinkel et al. 1998；Albertson et al. 2000；Bruder et al. 2001）是建立在杂交信号的强度随着阵列样点中 DNA 的复杂性增加而增加这一实验观测的基础之上。因为 BAC 阵列提供足够的精度以检测仅仅影响一部分克隆的单拷贝序列的变化，所以重叠 BAC 的阵列能够提供低达至少为 50 kb 的分辨率（Albertson et al. 2000）。

表 6-1　不同阵列 CGH 平台的比较

平台	阵列单元	单元复杂性	样品复杂性	描　述	优/缺点	参考文献
大插入片段基因组克隆	BAC、PAC、P1、黏粒	约 40 ~ 150 kb	高	阵列由分离自单个克隆的 DNA 点样制备	由于点样 DNA 具有高度复杂性，杂交信号很强。由于人和鼠基因组序列分析的不断进步，作图的 BAC 越来越容易获得。然而，从插入大片段的克隆载体中制备足量 DNA 很困难，且克隆的点样需要将长的 DNA 分子片段化	Solinas-Toldo et al.（1997）；Pinkel et al.（1998）；Albertson et al.（2000）
大插入片段基因组克隆的代表物	代表基因组克隆如 BAC、PAC、P1、黏粒的片段集合	通过产生代表物的方法降低克隆的初始序列复杂性	高	连接 介导 PCR 或 DOP-PCR 用于产生基因组克隆代表物	只需要少量从克隆分离的 DNA。连接-介导 PCR 保持了亲本克隆的复杂性，因此能产生和使用整个克隆一样的结果。如果代表物明显地降低样点中材料的复杂性，信号强度会下降	Snijders et al.（2001，Ligation-mediated PCR）；Hodgson et al.（2000，DOP-PCR）
cDNA 微阵列	PCR 扩增的 cDNA 和 EST	0.5 ~ 2 kb	高	用 PCR 扩增的 cDNA 点样制备阵列	使用商品化文库时，建立良好的点样方法可以实现 DNA 拷贝数和基因表达的平行测定 cDNA/EST 文库是不完整的，并且当前不是所有的 EST 均被精确作图。可能需要对数个在染色体上相邻的阵列单元的比率进行平均以精确鉴定单拷贝增益和丢失的区域。检测杂合性丢失的灵敏性较 BAC 阵列要低	Pollack et al.（1999）
Inter-*Alu* PCR 阵列	Inter-*Alu* PCR 片段	0.2 ~ 2 kb	降低	Inter-*Alu* PCR 扩增已作图的基因组克隆。单个片段被克隆并用于阵列制备。实验和参考基因组 DNA 可以相似地扩增，标记和与阵列杂交。	简化了从基因组克隆制备 DNA，并且要分析的 DNA 被显著富集。检测单拷贝的改变需要对数个相邻的阵列单元进行平均	Geschwind et al.（1998）
基因组代表阵列	代表性 PCR 片段	0.2 ~ 2 kb	降低	限制性内切核酸酶消化的基因组 DNA 与引物连接并扩增。单个扩增片段用作为阵列单元 阵列的复杂性比原始基因组 DNA 低 2 ~ 40 倍。标记的实验和参考 DNA 的代表物用相同的限制性内切核酸酶制备并与阵列杂交	代表物的制备提供了显著的扩增，因此只需要少量 DNA 进行分析 阵列单元的作图位置必须已经建立。实验和参考 DNA 必须来自同一人，除非需要进行多态性检测	Lucito et al.（1998，2000）

在第三种方法中，借助于 inter-*Alu* PCR（Geschwind et al. 1998）或者使用限制性酶和 PCR 生成基因组的代表物，即代表性差异分析（representational difference analysis，RDA）方法（Lucito et al. 2000）（方案 4 和方案 5），而显著地降低总基因组 DNA 的复杂性。制备微阵列时，所使用的 DNA 片段克隆自用同样方法生成的序列减少了的基因组代表物，标记的基因组代表物用来和该阵列进行杂交。

表 6-1 总结了目前能够使用的三种阵列 CGH 方法的性能和不足。它们之中没有任何一个在这一领域中占据优势地位，并且每种方法可能在特定的情况下有优势。对原始文献进行评估能够了解它们是否适合于特定的研究目的。

363

方案 1

cDNA 微阵列比较基因组杂交

Jonathan R. Pollack（斯坦福大学医学院，斯坦福，加利福尼亚 94305）

cDNA 微阵列为阵列 CGH 分析提供了一种有吸引力的平台，这部分归因于该技术的熟悉和强力。数年以来，包含上万个 cDNA 的高密度 cDNA 微阵列已经被常规性点样制备并用于表达分析（见第 1 章）。这些同样的阵列能够用于 CGH 分析，并且它们为基因组 DNA 拷贝数变异的比率测定提供了非常高的作图分辨率，以至达到逐个基因和外显子的水平。而且，应用相同的微阵列进行 mRNA 表达谱的平行分析，提供了随时在全基因组及靶染色体区域上，将 DNA 拷贝数与基因表达变异进行比较和联系起来的机会。这为将候选癌基因进行优先排序，将其用于更深入的研究提供了基础，例如，在扩增子中，预期癌基因在扩增时会高度表达。

cDNA 微阵列已经被广泛地用于确定癌细胞中 mRNA 的水平（DeRisi et al. 1996；Alizadeh et al. 2000；Perou et al. 2000）。因为基因组 DNA 的序列至少比哺乳动物细胞 mRNA 群体的序列空间在复杂性上高一个数量级，因此使用 cDNA 微阵列来测定基因组 DNA 拷贝数变异需要严格的方案，以达到所必需的灵敏度和精度。

这里的详细方案介绍了基因组探针的制备，及其与 cDNA 微阵列杂交用于高分辨率的 CGH 分析（图 6-1）。这些方法表明了使用 cDNA 微阵列的 CGH 有足够的灵敏性和高的精度（变异系数约 10%）检测基因组 DNA 序列的单拷贝丢失（Pollsck et al. 1999）。方案建立和数据分析的特别注意事项请参考方案最后部分**疑难解答**的信息栏。

364

材料

注意：标记有“〈!〉”的材料的正确处理方法请参见附录 3。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成请参见附录 2。

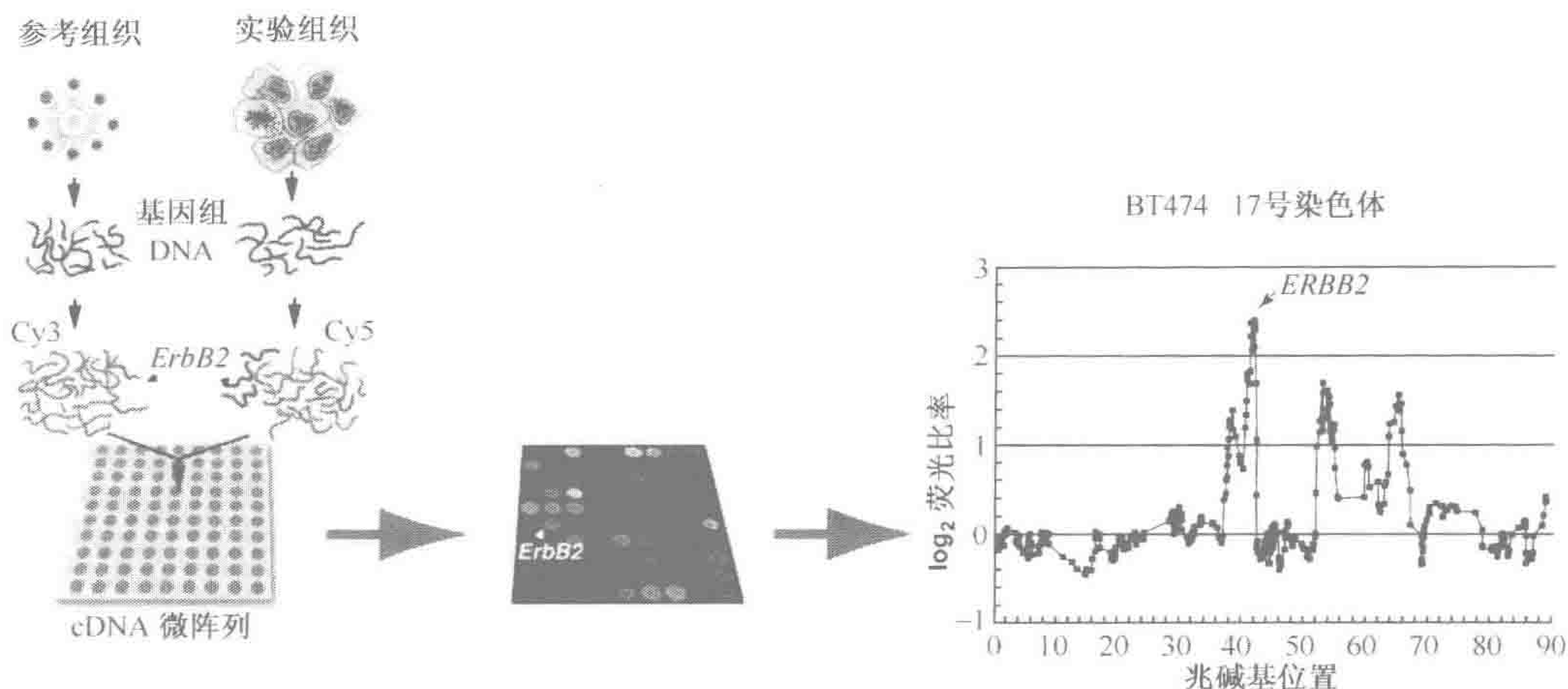


图 6-1 cDNA 微阵列比较基因组杂交示意图

实验和参考基因组 DNA 用两种荧光基团标记，并同时与 cDNA 微阵列杂交。对于阵列上每种 cDNA，两种荧光基团强度的比率反映了该基因在基因组中的相对拷贝数。荧光比率可以对染色体图谱位置进行作图，以得到实验样品中 DNA 拷贝数变异的高分辨率图谱。这里显示的结果描述了人乳腺癌细胞系 BT474 的 17 号染色体上 DNA 拷贝数的变异（经允许修改自 Pollack et al. 1999）

使用前将储存液稀释到合适的浓度。

EDTA（0.5 mol/L，pH 8.0）

2.5×随机引物混合物

125 mmol/L Tris-Cl（pH 6.8）

12.5 mmol/L $MgCl_2$

25 mmol/L β-巯基乙醇〈!〉

750 μg/ml 随机八聚体

这些成分可以单独配制，或者完整缓冲液可以作为 BioPrime DNA Labeling System 试剂盒（18094-011，Invitrogen）的一部分购买到。

SDS（10 %）〈!〉

20× SSC

TE（pH 7.4）

酶和缓冲液

E. coli DNA 聚合酶 I Klenow 片段（高浓度，40～50 单位/μl）

使用高浓度 Klenow 片段显著改善标记效率。该酶可以作为 BioPrime DNA Labeling System 试剂盒（18094-011 Invitrogen）的一部分购买到，或者从 New England Biolabs（M0210M）单独购买。

核酸和寡核苷酸

cDNA 微阵列

cDNA 微阵列，点样在聚-L-赖氨酸包被的显微镜载玻片上，可以自行制作（见第 2 章，方案 2）或购买商品化的产品。微阵列应在杂交前处理，用琥珀酸酐封闭未结合的聚-L-

赖氨酸（PART Ⅱ，方案 2）。使用其他非聚-L-赖氨酸材料修饰的玻璃表面可能导致荧光背景增加（见本方案最后疑难解答的信息栏）。

Cy5 和 Cy3 标记 dCTP（1 mmol/L 储存液）（PA55021 和 PA53021 Amersham Biosciences）dNTP-标记溶液：含有 dATP、dGTP、dTTP（每种的浓度为 1.2 mmol/L）和 dCTP [浓度为 0.6 mmol/L 的水溶液或 TE（pH 8.0）溶液]

使用超纯 dNTP（如 27-2035-01，Amersham Biosciences）。有关 dNTP 储存液的制备及保存的建议，请见附录 2。

BioPrime DNA Labeling System 试剂盒（Invitrogen）提供的 dNTP 标记混合物中包含生物素化的 dCTP，不能用于本方案中。

人 C_0t-1 DNA（1μg/μl 储存液）（15401-011，Invitrogen）

Poly（dA）：Poly（dT）（P 9764，Sigma）

用 TE（pH7.4）配制成 5 mg/ml 储存液。

已知浓度的参考 DNA

从正常男性或正常女性外周血白细胞制备的基因组 DNA，分别适于作为男性或女性实验样品的参考基因组 DNA。参考 DNA 和测试 DNA 应当用相同的方法制备。

已知浓度的实验基因组 DNA

使用任何标准方法制备的基因组 DNA 的质量都足以用于阵列 CGH 分析（见附加方案）。QIAGEN Blood & Tissue Genomic DNA 试剂盒（13362）提供了一种便利的选择。基因组 DNA 应当以 200～400 μg/ml 的浓度溶解在 TE 中。对于高复杂性的 DNA（如人基因组 DNA），如果首先用限制性内切核酸酶如 *Dpn*II 消化以减小 DNA 的大小，标记反应会进行得更加有效。消化后，DNA 应当用酚：氯仿抽提和乙醇沉淀来纯化。或者可以使用 QIAGEN PCR Purification 试剂盒（28106，QIAGEN）。

酵母tRNA（15401-11，Invitrogen）

用 TE（pH 7.4）配制成 5 mg/ml 储存液。

专用设备

带有微孔板适配器的台式离心机（如 Beckman Coulter Allegro 6R 或与之相当的）

荧光扫描仪（双色，如 Axon GenePix 3000）

玻璃盖玻片（22 mm×60 mm）（12-544-G，Fisher Scientific）

玻璃洗涤缸（3）（900303，Wheaton）

加热槽，预热至 100℃

杂交盒（510-459-9101，Monterey Industries，Richmond，California）

金属玻片架（2）（900234，Wheaton）

Microcon YM-30 过滤器（42411，Millipore）

水浴，预热至 37℃和 65℃

65℃水浴空间必须足够大，以便杂交和洗涤（如 Hot Shaker，Bellco Glass Co.）

方法

基因组 DNA 随机引物标记

1. 标记两个带旋盖的微量离心管，一个用于实验基因组 DNA，另一个用于参照

样品。

2. 向合适的管内加入 2～4 μg 要标记的实验和参考样品的基因组 DNA。用 dH_2O 将每支试管的体积调节至 21 μl。

   如果可能，最好使用 4 μg 基因组 DNA 作为起始材料，因为它能够增加标记产物的数量。在步骤 4 中加入额外的 Klenow 片段或在步骤 5 中延长孵育时间能够使较少量的 DNA 材料充分标记。

3. 向每支试管中加入 20 μl 2.5×随机引物混合物。将试管在加热槽中于 100℃下放置 5 min，然后快速转移到冰上。

   冷却后，如果需要，进行短暂离心使内容物集中在试管底部。

4. 向每份反应混合物（冰浴中）中加入：

dNTP 标记溶液	5 μl
Cy5 或 Cy3 标记的 dCTP（1 mmol/L）	3 μl
高浓度 Klenow 片段（40 单位/μl）	1 μl

   通常，实验样品的基因组 DNA 用 Cy5 标记，参考样品的基因组 DNA 用 Cy3 标记。Cy-标记 dUTP 可以替代 Cy-标记 dCTP，但是，dNTP 标记混合物中的核苷酸浓度必须做相应调整。

5. 轻敲试管壁混匀组分。将反应混合液在 37℃下温育 2 h，然后加入 5 μl 0.5mol/L EDTA（pH 8.0）终止反应。

6. 使用 Microcon YM-30 过滤器纯化标记的基因组 DNA。

   a. 将标记的实验和参考样品，转移到一个 Microcon YM-30 过滤器中。加入 450 μl TE（pH 7.4）。

   b. 将过滤装置放在微型离心机中，于室温下以 12 000 g 离心 9 min。继续离心额外的 30 s 到 1 min，直到标记 DNA 的体积为 15～32 μl。弃去流出液体。

   c. 向过滤器中另外加入 500 μl TE（pH 8.0），并重复离心，直至标记 DNA 的体积重新达到 15～32 μl。弃去流出液体。

7. 将下列成分转移到 Microcon YM-30 过滤器中：

TE（pH 7.4）	450 μl
poly（dA）-poly（dT）	20 μg
酵母 tRNA	100 μg
C_0t-1 DNA（人或鼠，根据需要）	50 μg

   进行细胞遗传学 CGH 时，标记的基因组 DNA 样品必须与过剩的合适 C_0t-1 DNA 预复性。预复性步骤用于封闭标记的高度重复序列与 cDNA 微阵列中相似序列的复性，如在约 5% 的阵列 cDNA 克隆的 3′ 端非翻译区域（3′ UTR）出现的重复 *Alu* 元件。重复元件的封闭失败会导致荧光比率的压缩。我们的经验还发现必须用过剩的 poly（dA）-poly（dT）与标记的基因组 DNA 样品进行预复性，以封闭与阵列中 cDNA 的 3′ 端尾巴杂交（见本方案最后**疑难解答**的信息栏）。

8. 按照步骤 6 的方法离心样品（约 11 min），直至标记样品的体积约 15～32 μl。倒转过滤器并将其放置在一支新的微量离心管中。将试管在室温下以 12 000 g 离心 1 min 以回收标记的 DNA。

杂交

9. 将标记探针转移到新的带旋盖的微量离心管，并用水将体积调节到 32 μl。加入

6.8 μl 20×SSC（最终浓度为 3.4×SSC），并轻敲管壁混合试剂。加入 1.2 μl 10% SDS（最终浓度为 0.3 %），并再次轻敲管壁混合试剂。

杂交混合物的体积（40 μl）足够用于在 22 mm×60 mm 玻璃盖玻片下方与 cDNA 微阵列复性。如果使用其他规格的盖玻片，对杂交混合物体积作相应调整。

10. 将试管在加热槽中于 100℃下放置 2 min，然后将试管在 37℃下温育 25 min。

    37℃下温育使得在杂交混合物转移到 cDNA 微阵列上之前，poly（dA）-poly（dT）和 C_0t-1 DNA 与标记的基因组 DNA 复性。

11. 室温下，将样品在微型离心机中以最大速度离心 5 min。

    本步骤是可选的，但是它能减少微阵列表面上荧光碎片的数量。

12. 将 cDNA 微阵列放入杂交室中。将 40 μl 杂交混合物加到 cDNA 微阵列上，然后仔细盖上玻璃盖玻片（22 mm×60 mm）。为了保持湿度，向杂交室的储液池中加入 20 μl 3×SSC。如果没有储液池，将 3× SSC 在盖玻片表面分散为约 1 μl 的液滴。

    微阵列玻片的操作，见第 2 章。

13. 封闭杂交室，将其放在 65℃水浴中。将微阵列杂交反应在 65℃下孵育 16～20 h。

洗涤和扫描

14. 在分开的玻璃洗涤缸中准备下面一系列洗涤溶液，各 400 ml：

    2× SSC/0.3% SDS（溶液预热到 65℃）

    2× SSC（溶液预热到 65℃）

    1× SSC

    0.2× SSC

    准备第一种洗涤液时，先用水稀释适量的 20× SSC 储存液，然后加入适量 10% SDS 储存液，否则 SDS 会从溶液中沉淀出来。

15. 将杂交室从水浴中取出，用吸水纸擦干外周。将微阵列从杂交室中取出，立即放置到第一种洗涤液中的金属玻片架上。根据下列洗涤程序，将玻片架在不同玻璃缸中转移，并轻微振荡来洗涤微阵列。

	SSC	SDS	温度	时间
洗涤 1A：	2×	0.3%	65℃	5 min
洗涤 1B：	2×	—	65℃	短时漂洗
洗涤 2：	1×	—	20℃	5 min
洗涤 3：	0.2×	—	20℃	5 min

洗涤 1A 中，小心地让盖玻片从微阵列上轻轻滑落，以减少擦伤阵列表面的可能性。第一次洗涤在 65℃下进行满 5 min 看来对增加特异性和非特异杂交信号的比率显得非常重要。在洗涤 1B 中，在 2×SSC 中短时清洗微阵列，然后将它们放在新的玻片架上进行洗涤 2，以防止带入 SDS，它是自发荧光来源之一。

16. 将金属玻片架放入带有微量滴定板适配器的台式离心机中。室温下以 46 *g*（500 r/min）离心 2 min 甩干微阵列。

将微阵列离心甩干对于减少在空气中干燥时盐沉淀引起的荧光背景很重要。

17. 用双色荧光扫描仪对 cDNA 微阵列进行扫描。

建议在洗涤后数小时内对杂交的微阵列进行扫描，因为荧光信号经过数小时到数天后将衰减（特别是在 Cy5 通道）。将微阵列保存在密闭的玻片盒中。

数据分析

- **标准化：**为了补偿实验和参考 DNA 起始浓度上的差别，以及实验（Cy5）和参考（Cy3）DNA 标记效率上的差异，必须对阵列上所有单元的荧光比率进行标准化。标准化确保了 DNA 拷贝数没有变化的位点，相对于阵列上所有单元的平均比率测量值，荧光比率（实验/参考）为 1。

  通过将所有荧光比率乘以一个因子，使得阵列上所有测量良好的基因的平均荧光比率为 1，就实现了比率测量的归一化（这一计算在对数空间进行，以便于对 DNA 扩增和缺失同等地加权）。注意：这种标准化方法和大多数 mRNA 水平的微阵列分析中的常规处理方法相同（见第 1 章，PART Ⅲ）。对于细胞遗传学 CGH 而言，这种标准化方法确保荧光比率反映了相对于实验样品中增益/丢失平均水平的 DNA 增益和丢失。但是，例如，它不能区分如一个纯四倍体 DNA 样品和二倍体 DNA 参照。

- **荧光比率在基因组上作图：**阵列上单个的 cDNA 可以通过基因名称加以识别，对于匿名 EST，则用 UniGene 聚类（Schuler 1997）识别。直到最近，利用辐射杂交系（radiation hybrid，RH）作图，得到了分辨率最高的人类基因组上 cDNA 图谱（Deloukas et al. 1998）。随着人类基因组草图序列的发布（Lander et al. 2001），阵列中的 cDNA 现在可以有更高的精度按照局部顺序在人类基因组上作图。为了产生染色体中 DNA 拷贝数（荧光比率）的变异图谱（图 6-1），必须将阵列单元中 cDNA 存取/基因名称和染色体的核苷酸作图位置关联起来。数据库如 University of California，Santa Cruz（http://geonme. ucsc. edu/）维护的有助于这一目标。也可以得到进行这项任务的基于 Web 工具形式的脚本（如 CaryoScope；http://genome-www. stanford. edu / ～rees；C. Rees and J. Pollack，未发表）。有关成像和数据分析的更多详细内容，见 Pollack et al.（1999）。

疑难解答

- **起始：**在第一次建立分析时，研究人员可能会发现与已知拷贝数变化的基因组 DNA 杂交是非常有用的，例如，通过细胞遗传学 CGH 已经很好地确定了 DNA 拷贝数变化的癌细胞系，或者含有不同数目 X 染色体的细胞系（如 49，XXXXX 与 46，XX 相比较），对于所有定位在 X 染色体上的阵列 cDNA，后者模拟了不同水平的增益或丢失。这些基因组 DNA 试剂中的许多可从 ATCC（www. atcc. org）或者 NIGMS 人类遗传突变（human genetic mutant）细胞库（locus. umdnj. edu/nigms/）获得。对于 RNA 杂交，从基因组 DNA 杂交中得到的荧光比率是成线性比例的，但是倾向于低估拷贝数的变化（Pollack et al. 1999）。

  自身杂交（如 46，XX 与自身相比较）可用于确定荧光比率的范围及标准偏差。该信息有助于设定阈值，高于或低于它的异常比率被认为反映了真正的 DNA 变化。

  必须永远包含在微阵列中作为对照样点的有：人基因组总 DNA、人 C_0t-1（用于评价封闭的效率）、酵母 DNA 和 3×SSC（评价杂交特异性）。

- **低信噪比：**研究人员可以预期平均信噪比为 5，这个信噪比足以进行高精度测定（变异系数约 10%）。据我们了解，使用随机引物标记基因组 DNA 要比切口平移标记取得的结果更好。标记前减小基因组 DNA 的大小和使用高浓度 Klenow 酶，对于取得最佳标记和强杂交信号来说都很

重要。尽管标记效率不容易定量，强烈标记应产生深紫色杂交混合物（颜色比一般 mRNA 标记反应中见到的要明显更深）。

高背景可导致低信噪比，例如，由于使用其他的阵列表面（本方案是为聚-L-赖氨酸包被的玻片表面优化）或聚-L-赖氨酸包被玻片处理不充分（见第 2 章，PART Ⅰ）导致信噪比降低。为了区分基因组 DNA 标记问题和微阵列玻片问题，进行 mRNA 样品的标记和杂交是有用的，预期这样会产生高背景。

- **强信号，但荧光比率被压缩**：如果杂交信号强，但荧光比率（实验：参考）显著低估了真实比率，则说明杂交的特异性不是最佳。与本方案中详细介绍的使用水相杂交液相比较，使用基于甲酰胺的杂交液可能导致杂交特异性降低。第一次杂交后洗涤在 65℃ 下进行，对获得高的杂交特异性显得非常重要。
- **在大多数实验（肿瘤）样品中观察到有高荧光比率的 cDNA 子阵列**：基因组 DNA 肿瘤样品看来相对富集了 poly（dA）-poly（dT）序列。因此在杂交混合物中必须包含 poly（dA）-poly（dT），以封闭这些标记序列与 cDNA 阵列单元中含有长的 poly（dA）尾巴的子阵列发生复性。封闭不充分可能会造成可与这些带有 poly（dA）尾巴的 cDNA 复性的基因组序列在大多数肿瘤样品中表现出被扩增了。

  在 cDNA 微阵列制备的早期，cDNA 克隆没有进行必要的克隆纯化或序列验证。这导致在一小部分阵列单元中包含了线粒体 DNA 污染物。因为线粒体 DNA 通常在肿瘤细胞中普遍存在，在大多数肿瘤样品中，这些单元的荧光比率表现出增高。虽然这在较新的阵列中并不是一个显著的问题，通过标记的人线粒体 DNA 与 cDNA 微阵列杂交能够容易地鉴别出线粒体 DNA 污染。
- **某些 cDNA 荧光比率不能反映预期的拷贝数变异**：作者观察到在他们的微阵列上，一小部分 cDNA 单元不能准确地测定预期的 DNA 拷贝数变异。这些发现部分反映了 cDNA 与定位在多条染色体上的基因（基因家族，假基因）存在高度相似性，以及 cDNA 没有被正确地鉴定（序列-追踪错误），或者是在人基因组上被错误地定位。因此，使用荧光比率的“移动平均（moving average）”（Pollack et al. 1999），将染色体上相邻的阵列单元进行平均，可能得到染色体上 DNA 拷贝数变异更为精确的图谱。

370

方案 2

制备 BAC DNA 用于通过连接介导 PCR 进行 CGH

Antonie Snijders,*† **Daniel Pinkel**†‡，**Dorrna Albertson***†‡ *综合性癌症中心、†癌症研究所、‡实验医学系，加州大学，旧金山，加利福尼亚，94143

可获得的已作图 BAC（细菌人工染色体）克隆集的数量正日益增多（Cheung et al. 2001），这极大地简化了对含有 DNA 拷贝数变化的基因组区域的鉴定。然而，制备 BAC DNA 用于微阵列点样却是一个挑战性的工作。因为 BAC 是单拷贝载体，从细菌培养物中制备 DNA 的产量要比从含有高拷贝数载体培养物中制备的产量低。而且，当使用足够的浓度点样黏性的高分子量 DNA，以产生高信噪比时，难以获得实验的一致性（Snijders et al. 2001）。为了从克隆在高容量载体中的插入片段制备足够的基因组 DNA 以用于制作阵列，采用了许多方法来扩增插入片段的代表，包括连接介导 PCR（Snijders et al. 2001）和简并寡核苷酸引物（DOP）—PCR（Hogdson et al. 2001）。连

接介导 PCR 有望更加完好地保留 BAC DNA 的复杂性，从而提供较好的阵列性能。

这里介绍的生成基因组克隆 DNA 的高复杂性代表的方法，其结合了连接介导 PCR 方法和随后的进一步扩增，以此克服将大插入片段克隆用于 DNA 微阵列 CGH 时遇到的限制（Klein et al. 1999；Snijders et al. 2001）。

图 6-2 和图 6-3 提供了方案步骤的概述。用限制性内切核酸酶 *Mse* Ⅰ 消化 BAC DNA，然后接头与消化的 BAC DNA 复性，使用连接介导 PCR 完成第一轮扩增（图 6-2）。接着使用少量（1 ng）扩增产物作为随后扩增的模板。第二轮扩增产物纯化后，重悬在 20%二甲基亚砜（DMSO）中，浓度为 0.8 μg/μl（点样液），用于阵列点样。整个操作过程从 1 ng BAC DNA 得到的点样液可以足够制作数万个阵列，实际上为阵列制备提供了不受数量限制的试剂。

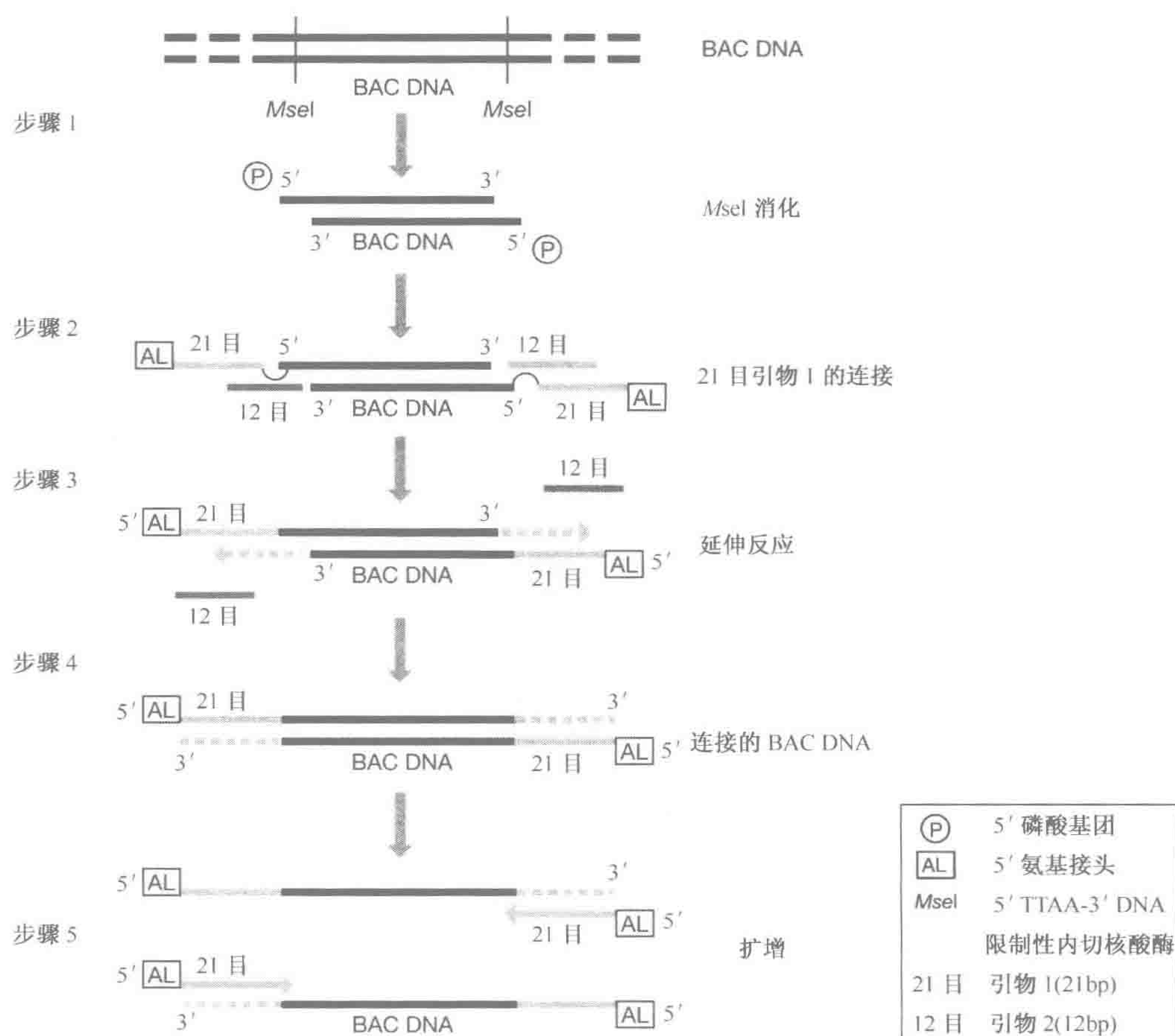


图 6-2 连接介导 PCR 示意图

BAC DNA 用限制性内切核酸酶 *Mse* Ⅰ 消化，产生一组携带 5′磷酸化突出末端的 DNA 片段（步骤 1）。引物 2 引导引物 1 到突出的序列，使得引物 1 与 BAC DNA 片段连接（步骤 2）。由于在其 5′端缺少磷酸基团，引物 2 与 DNA 片段的 3′端不能连接。加热处理从 DNA 片段 3′端脱去引物 2，然后在由 *Taq*I 和 *Pwo* DNA 聚合酶共同催化的合反应中，暴露出的 3′端延伸（虚线）（步骤 3 和 4）。然后用得到的 DNA 作为模板，由连接后剩余的过量引物 1 作为引物，进行扩增反应（步骤 5）。进行第二轮扩增以生成点样用的 DNA（未显示）（经允许，修改自 Snijders et al. 2002）

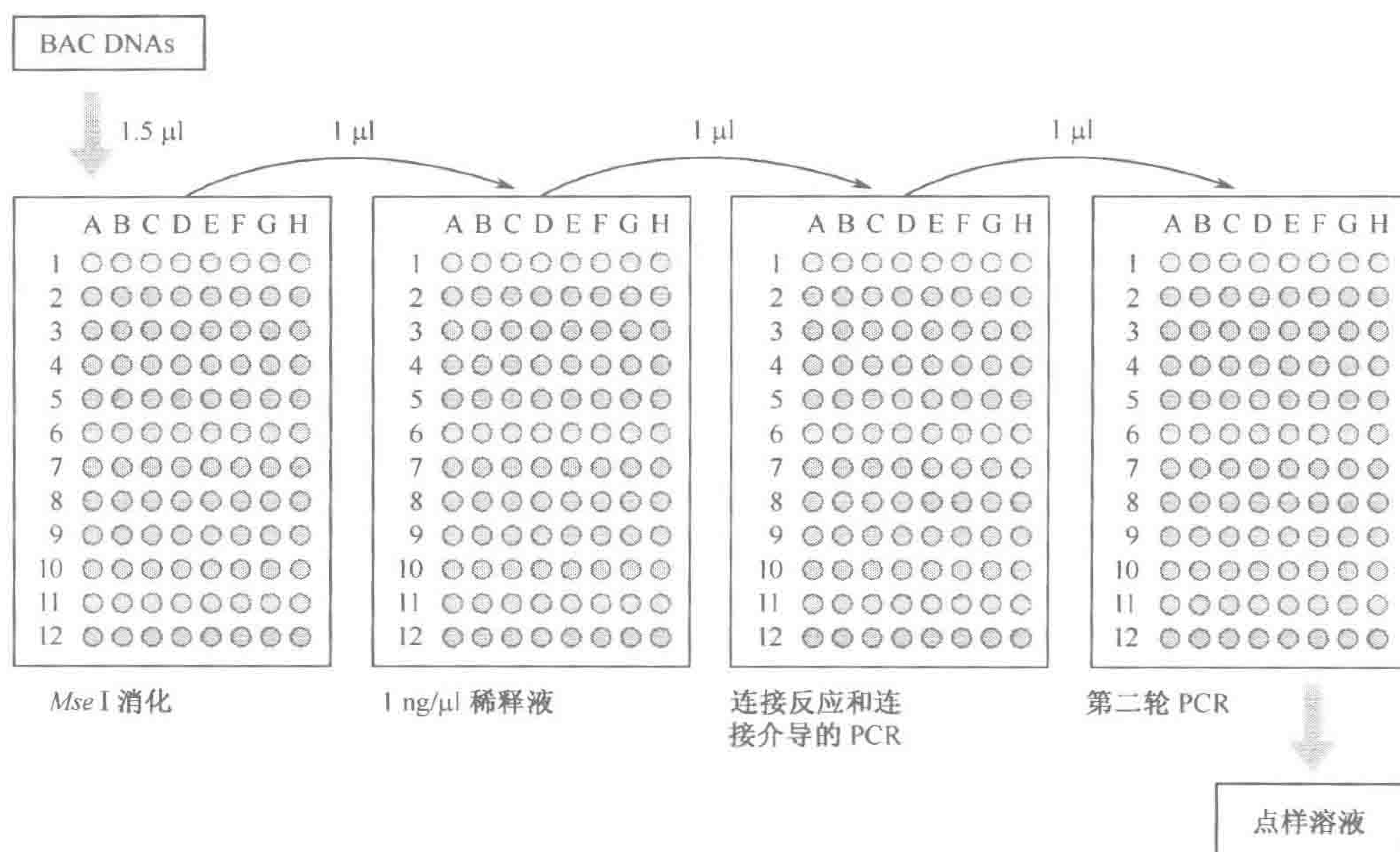


图 6-3　96 孔板上 BAC DNA 克隆的连接介导 PCR 方法示意图

BAC DNA 消化后，每个消化反应产物取 1 μl 用于制备浓度为 1 ng/μl 的 DNA 稀释液；取 1 μl 稀释液用于连接及随后的 PCR 反应。在该 PCR（50 μl）中，只使用 1 μl 产物作为第二轮 PCR 的模板。DNA 的最终产量约 10～15 μg，沉淀之后重新溶解于 12 μl 12% DMSO 中。将所得溶液加入微孔板用于点样。注意：从约 20 ng 的起始 BAC DNA 能获得约 1000 份点样溶液（起始 DNA 扩增约 5×10^5）（经允许，修改自 Snijders et al. 2002）

在重复微阵列杂交实验中（见方案 3），单独制备的 BAC 代表产生高度可重复性的数据：来自相同克隆的两个独立制备物线性比率的平均变异为 6.6%。而且，使用由给定 BAC 克隆集制备的代表物点样的阵列中测量得到的比率，和使用直接从相同的克隆分离未扩增的 BAC DNA 点样制备的阵列中得到的那些比率，实质上是相同的（Snijders et al. 2001）。

有关 BAC 克隆载体特性和 BAC 克隆纯化更多的信息，见 Sambrook 和 Russel（2001）书中第 4 章。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

ATP（10 mmol/L）（来自 Invitrogen）

　　在 −20℃ 下保存。

DMSO（20%）溶于灭菌 H_2O

溶液在室温下保存。

dNTP 溶液，含有全部四种 dNTP，每种核苷酸浓度都为 10 mmol/L 和 25 mmol/L

乙醇（无水、冰冻和 70%）

有报道称，有些批次的乙醇（100%或 200 标准强度）中含有会使微阵列试验中产生很深荧光背景的污染物质，95%的乙醇不存在这个问题。为避免这个问题，可以用 95%的乙醇或者用已鉴定证明不会产生高荧光背景的无水乙醇。Warner-Graham 公司（Cockeysville，Maryland）出售的无水乙醇（USP 级乙醇，64-17-5）效果很好。

H_2O（灭菌，如高压灭菌、去离子和过滤）

$MgCl_2$（25 mmol/L）（Perkin Elmer）

在－20℃下保存。

10× One-Phor-All Buffer Plus（27-0901-02，Amersham Biosciences）

100 mmol/L Tris-乙酸（pH 7.5）

100 mmol/L 乙酸镁

500 mmol/L 乙酸钾

缓冲液在 4℃下保存。

One-Phor-All Buffer Plus 是一种通用型缓冲液，在本方案中以 2.2×用于 *Mse*I 消化，以 2.5×用于连接反应。

醋酸钠（3 mmol/L，pH 5.2）

TE（pH 7.4）

酶和缓冲液

373

AmpliTaq Gold DNA 聚合酶（5 单位/μl）（Perkin Elmer）

在－20℃保存。使用时将酶保持在冰上，用完后尽快将酶放回到－20℃。

DNA 聚合酶混合物，含有 *Taq* 和 *Pwo* DNA 聚合酶（3.5 单位/μl）

从 Roche 可获得长链模板 PCR 系统。

缓冲液在－20℃下保存。使用时将酶混合物保持在冰上，用完尽快将酶放回到－20℃。

Mse Ⅰ 限制性酶（20 单位/μl）（New England Biolabs）

酶在－20℃下保存。

10× PCR 缓冲液Ⅰ

缓冲液在－20℃下保存。

从 Roche 可获得长链模板 PCR 系统。

10×*Taq* 缓冲液Ⅱ（不含 $MgCl_2$）（Perkin Elmer）

缓冲液在－20℃下保存。

T4 DNA 连接酶（5 单位/μl）（Invitrogen）

连接酶在－20℃下保存。使用时将酶混合物保持在冰上，用完尽快将酶放回到－20℃。

凝胶

超纯（Ultrapure）琼脂糖（Invitrogen）制备的琼脂糖凝胶（1%），用 1×TBE 倒胶，胶中含 0.5 μg/ml 溴化乙啶〈!〉

6×上样缓冲液

0.25% 溴酚蓝〈!〉

0.25% 二甲苯氰〈!〉

30% 甘油水溶液

缓冲液 4℃储存。

核酸和寡核苷酸

BAC DNA（20～500 ng），经纯化将宿主 DNA 污染减至最低

BAC DNA 可以按照 Sambrook 和 Russel（2001）书中第四章介绍的方法进行纯化。

用于连接 BAC DNA 和扩增连接产物的引物

引物 1：5′-AGT GGG ATT CCG CAT GCT AGT-3′（估计要用 50 nmol/L，含有 5′氨基连接臂）

- 引物储存液：以 500 μmol/L 浓度溶于灭菌的 TE（pH 7.4）。
- 100 μmol/L 工作液：将储存液按 1∶5 用灭菌水稀释。固态引物、储存液、和工作液在－20℃下保存。

引物 2：5′-TAA CTA GCA TGC-3′（估计要用 50 nmol/L）

- 引物储存液：以 500 μmol/L 浓度溶于灭菌的 TE（pH 7.4）。
- 100 μmol/L 工作液：将储存液按 1∶5 用灭菌水稀释。固态引物、储存液、和工作液在－20℃下保存。

ϕX174 RF DNA/*Hae*Ⅲ 分子质量标记（Promega）

专用设备

可加热到 45℃的鼓风（杂交）烘箱（如 Techne Hybridiser HB-1D）

烘箱用于干燥 PCR 产物（步骤 31）。

多通道移液器[如一支移液体积为 0.5～12.5 μl 的 8 或 12 道移液器（Matrix Technologies）和一支移液体积为 5～250 μl 的 8 或 12 道移液器（Matrix Technologies）]

平板封口膜（96 孔）（USA Scientific）

带有分离式 8 连排盖的聚丙烯 PCR 8 连排管（0.2 ml）（USA Scientific）

连排管和盖都必须无 RNase、DNase、DNA、热原及 PCR 抑制剂。

带隔板的试剂槽（25 ml）（Matrix Technologies）

热循环仪，可每分钟调温约 1.3℃（如 Perkin Elmer 9700）

紫外透射仪（如 Stratagene Stratalinker 2400，可产生 260 000 μJ UV）

374

方法

重要事项：根据方案中提示使用多通道移液器。

Mse Ⅰ 消化 BAC DNA

1. 使用灭菌水稀释 10× One-Phor-All Buffer Plus 至最终浓度为 0.8×，配制 750 μl，

然后向 8 连排管的每管中加入 93 μl 稀释缓冲液。

限制酶的稀释缓冲液足够用于进行约 3×96 个反应体积为 5 μl 的消化实验，剩余的稀释液在盖好 8 连排盖后可以在 4℃下保存。

2. 使用多通道移液器向 12 个 8 连排管 PCR 管的每支试管中加入 2.5 μl 0.8×缓冲液，并用 96 孔板封口膜密封试管，以防止蒸发和可能的污染。

一次同时处理 96 份不同的 BAC DNA 时，DNA/PCR 污染是可能发生的。所以，始终要用 10 mmol/L HCl（或其他商品化试剂）擦洗移液器。如果微孔板用 96 孔板封口膜封闭，在移去密封前离心，以便除去封口膜周围的水分，并且要缓慢地除去封口膜。保持 8 连排试管条、8 连排盖子、试管以及其他实验室的器具洁净。在限定的空间，如单独的房间或者是 PCR 工作站内进行扩增反应。不要在限定的空间中打开装有 PCR 产物的试管。建议在整个方案中使用带滤芯的吸头。

3. 使用单道移液器向每支试管中加入 1.5 μl BAC DNA。在将 BAC DNA 加入每个 8 连排管中后，把试管条放在不同的架子上，并用 96 孔板封口膜密封。

BAC DNA 的浓度通常为 20～400 ng/μl。注意不论每种 BAC 制备物中 DNA 的浓度如何，每个限制性酶消化反应使用 1.5 μl 的 BAC DNA。

在本步骤中，每次仅处理一条 8 连排 PCR 管。剩下的 11 个 8 连排试管条仍然用 96 孔板封口膜封闭。

- 向后翻卷含有消化缓冲液溶液的 96 孔板（存放板）封口膜，直至第一个 8 连排试管条暴露出来。
- 取出第一个 8 连排试管条，并向每管中加入 BAC DNA。
- 将此 8 连排试管条放入一洁净的 96 孔板（接受板），并将封口膜覆盖在 96 孔板上方，使封口膜覆盖整个 96 孔板（封口膜一头盖住 8 连排试管条，另一头粘在 96 孔板上）。
- 返回到存放板，将封口膜向后翻卷，直至暴露出下一个 8 连排试管条。取出试管条并向每管中加入 BAC DNA。部分揭开接受板上的 96 孔板封口膜，同时仍旧保持第一个 8 连排试管条被覆盖着。
- 用一只手拿着 96 孔板封口膜的同时，将下一个 8 连排试管条放入接受板中。重新密封接受板并重复这一步骤，直至所有 8 连排试管条处理完毕。

4. 使用 10× One-Phor-All Buffer Plus 稀释 *Mse*I 限制性酶，配制 120 μl 最终浓度为 2 单位/μl 的酶溶液。该过程中将酶及稀释液保持在冰上。弃去任何未使用的酶稀释液。

5. 分别向每管中加入 1 μl（2 单位）稀释的 *Mse* Ⅰ 酶，该过程中将酶稀释液保持在冰上。加入稀释的酶后，每个 8 连排试管条用一个 8 连排盖子盖住，并将试管条置于冰上。

375

使用多通道移液器吸取小量（约 1μl）酶是不可靠的。作者发现使用单道移液器分别吸取酶，如 *Mse* Ⅰ、T4 DNA 连接酶和 DNA 聚合酶分解混合物（resolution mixture）会得到更为可靠的结果。

6. 将反应在热循环仪中，于 37℃下保持 12～16 h（或过夜）。

7. 用含有 0.5 μg/ml 溴化乙啶的 1%琼脂糖凝胶进行常规电泳，分析 1.75 μl 的小分量消化产物。加入 DNA 分子质量标记（如 ϕX174 RF DNA/*Hae* Ⅲ）以检查

片段长度。消化产物应为 100～1500 bp。

特异性引物与 BAC DNA 连接

连接反应在 10μl 反应体系中进行，包含 5 μmol/L 引物 1、5 μmol/L 引物 2、0.5× One-Phor-All Buffer Plus、1 mmol/L ATP、5 个单位的 T4 DNA 连接酶和 1 ng 消化的 BAC DNA。

8. 用灭菌水稀释 1 μl 每种消化产物，使 DNA 最终浓度为 1 ng/μl。
9. 准备 12 个新的 8 连排试管条。
   a. 向每一新管中分别加入经过估算体积的无菌水，其位置与每个消化反应产物的位置对应。
   b. 使用多通道移液器转移 1 μl 每种消化反应产物。
   c. 用 96 孔板封口膜覆盖（图 6-3）。
10. 在一支 1.5 ml 的离心管中混合下列成分（总体积 784 μl）以建立连接反应体系：

引物 1（100 μmol/L）	56 μl
引物 2（100 μmol/L）	56 μl
One-Phor-All Buffer Plus	56 μl
无菌水	616 μl

    反复吸打以混匀，并向一个 8 连排试管条的每管中加入 98 μl 引物溶液。
11. 用多通道移液器从 8 连排试管条的每管中吸取 8 μl 引物溶液（步骤 10），加入 7 μl 到 12 个新的 8 连排试管条的管底。用 96 孔板封口膜封住 8 连排试管条。

    使用一支 12.5 μl 多通道移液器吸取 8 μl 引物溶液，加入 7 μl，弃去剩余的 1 μl 及移液器吸头。为保证准确，鼓励对每一个 8 连排试管条使用新的移液器吸头。
12. 向每个对应 8 连排试管条中的 7 μl 引物溶液内加入 1 μl（步骤 9）制备好的 1 ng/μl BAC DNA（图 6-3）。
    a. 小心揭开 BAC DNA 稀释液微孔板的密封膜，然后用多通道移液器从第一个 8 连排试管条吸取 1 μl。
    b. 只启封第一个含引物溶液的 8 连排试管条，取出试管条，将 DNA 直接加入 7 μl 引物溶液中，加入 DNA 后将两个 8 连排试管条封闭。
13. 将 12 个 8 连排试管条转移到热循环仪中。复性反应在 65℃下进行 1 min，然后以每分钟降低约 1.3℃降温到 15℃（如 Perkin Elmer 9700 热循环仪中 5%的斜率）。
14. 在复性过程中，向 8 连排试管的每管中加入 62.5 μl 10mmol/L ATP 溶液。盖上 8 连排试管并置于冰上。

    ATP 溶液足够进行约 5×96 个连接反应，剩余 ATP 溶液用 8 连排盖子盖好后可在 4℃保存。
15. 当热循环仪达到 15℃时，从机器中迅速取出管子，并小心打开所有的 8 连排试管条，包括含 ATP 溶液的 8 连排试管条（步骤 14）。使用设定为重复-吸液-模式的多通道移液器吸取 12.5 μl 10 mmol/L ATP，并向 96 支含有 DNA 试管的

每支试管内壁上加 1 μl。轻弹管架使 ATP 滑进 DNA/引物溶液中。用 96 孔板封口膜封住管子。

以这种方式加入 ATP 减少了移液中可能的污染。

16. 分别向 96 支试管的每支中加入 1 μl（5 单位）T4 DNA 连接酶。酶加入之后，用 8 连排盖子盖住 8 连排试管条，并置于冰上。
17. 将反应放在热循环仪中，于 15℃下保持 12～16 h（或过夜）。

连接介导 PCR

18. 在一支 15ml 试管中混合下列成分，总体积为 4480 μl：

10× PCR 缓冲液 1	336 μl
10 mmol/L dNTP 溶液	224 μl
无菌水	3920 μl

短暂涡旋混匀，并将稀释的 dNTP 溶液放置在冰上。将溶液倒入试剂槽中。

每个连接介导 PCR 在 50 μl 体系中完成，这个体系中包含 0.6× PCR 缓冲液 1、0.4 mmol/L dNTP 混合物、3.5 单位 DNA 聚合酶混合物及 10 μl 连接混合物。

19. 从热循环仪中取出连接反应（步骤 7），并小心打开所有 8 连排试管条。
20. 使用设定为重复-吸液-模式的多通道移液器吸取 205 μl 稀释的 dNTP 混合物，并向 5 个 8 连排 PCR 管的每管内壁上加 40 μl。剩余 5 μl 稀释的 PCR 混合物排回槽中，并弃去移液器吸头。
21. 重复该过程，将稀释的 dNTP 溶液转移到下一组的 5 个 8 连排 PCR 管中。每次操作完成时，将剩余 5 μl 稀释的 PCR 混合物排回槽中。
22. 调节多通道移液器使吸液体积为 85 μl，保留移液体积为 40 μl。从槽中吸取 85 μl 稀释的 dNTP 混合物，并将其中 40 μl 加入到剩余的 2 个 8 连排 PCR 管的内壁。轻弹 PCR 架使稀释的 dNTP 混合物滑进 DNA/引物溶液中。用 8 连排盖子密封试管。
23. 为了将引物 2 从 BAC DNA 上解离，将 PCR 管置于热循环仪中于 68℃下保持 4 min。然后取出试管，并向每管中加入 1 μl DNA 聚合酶混合物（3.5 单位/μl）。每次打开一条 8 连排试管，加入酶混合物后立即盖好，并置于冰上。

68℃温育 4 min 脱去引物 2（12 mer，T_m＝约 40℃）（图 6-2）。快速向每个 8 连排试管条中加入 DNA 聚合酶混合物，并将试管条置于冰上，这一点非常关键。本步骤中操作迅速对防止引物 2 显著复性很重要。随后的 68℃下 3 min 温育将除去任何重复性的引物 2，使得 BAC DNA 的游离 3′ 端可被延伸（图 6-2）。

从热循环仪中移出 12 个 8 连排试管条分别向每管中加入酶混合物、将 8 连排试管条置于冰上，以及最后将 12 个 8 连排试管条放回到热循环仪中进行扩增，整个过程应该不超过 10～15 min。

24. 将 8 连排试管条放到热循环仪中，并按下列变性、复性和聚合时间及温度进行扩增反应：

循环次数	变性	复性	聚合
第一次循环			68℃ 3 min
随后的循环（14）	94℃ 40 s	57℃ 30 s	68℃ 1 min，15s
随后的循环（34）	94℃ 40 s	57℃ 30 s	68℃ 1 min，45 s
最终循环	94℃ 40 s	57℃ 30 s	68℃ 5 min

扩增完成后，将 8 连排试管条置于 4℃下。

25. 每个扩增反应取 3.5 μl，用含有 0.5 μg/ml 溴化乙啶的 1%琼脂糖凝胶进行常规电泳分析。使用 DNA 分子质量标记（如 ϕX174 RF DNA/*Hae*Ⅲ分子质量标记）检查片段长度。扩增产物大小为 70～1500 bp，浓度最高的扩增产物约为 200～800 bp。

    如果带型显得模糊，说明连接和（或）连接介导 PCR 不成功，应该重复实验。所有产物的扩散应该为 70～1500 bp。任何偏离此范围的扩散，如约 300～600 bp 或者高分子质量的扩散都是不能接受的。

连接介导 PCR 产物的扩增

以连接介导 PCR 的产物作为模板，在常规扩增反应（Re-PCR）中生成微阵列点样用的 DNA。扩增在 100 μl 反应体系中进行，含有 1× *Taq* 缓冲液Ⅱ、引物 1（4 μmol/L）、dNTP 混合物（0.2 mmol/L）、$MgCl_2$（5.5 mmol/L）、2.5 单位 AmpliTaq Gold、和连接介导 PCR 产物（1 μl）。

26. 在一支 15ml 试管中混合下列试剂准备 PCR 混合物，总体积为 9900 μl：

10× *Taq* 缓冲液 Ⅱ	1000 μl
引物Ⅰ（100 μmol/L）	400 μl
dNTP 溶液（25 mmol/L）	80 μl
$MgCl_2$（25 mmol/L）	2200 μl
AmpliTaq Gold（5 单位/μl）	50 μl
无菌水	6170 μl

27. 将 PCR 混合物（步骤 26）倒入试剂槽中。用多道移液器向 12 个 8 连排试管条的每管中加入 99 μl 混合物。用 96 微孔板封口膜密封试管。
28. 用多道移液器按照下述步骤向每个对应的 PCR 混合物 8 连排试管条中加入 1 μl 连接介导 PCR 产物：
    a. 将一个 PCR 混合物 8 连排试管条（步骤 27）和一个装有连接介导 PCR 产物的 8 连排试管条（步骤 24）的密封膜揭开。
    b. 将 1 μl 模板（连接介导 PCR 产物）加入到 PCR 混合物中，盖上 8 连排试管条，然后再盖上装有剩余的连接介导 PCR 产物的 8 连排试管条。
    c. 按照步骤 a 和 b，继续下一组 8 连排试管条，直至所有模板均处理完毕。
29. 将含有 100 μl PCR 混合物和 1 μl 连接介导 PCR 产物的 8 连排试管条放入到热循环仪中，并按照下列变性、复性和聚合时间及温度进行扩增反应：

循环次数	变性	复性	聚合
第一次循环	95℃ 10min		
后续循环（45）	95℃ 30 s	50℃ 30 s	72℃ 2 min
最终循环			72℃ 7 min

扩增完成后，将 8 连排试管条于 4℃下温育。

30. 每个扩增反应取 4 μl，用含有 0.5 μg/ml 溴化乙啶的 1%琼脂糖凝胶进行常规电泳分析。使用 DNA 分子质量标记（如 ϕX174 RF DNA/HaeⅢ分子质量标记）检查片段长度。扩增产物大小应为 200～1500 bp。

通过 Re-PCR 制备点样溶液用于阵列 CGH

31. 干燥样品使扩增产物体积减少到约 50 μl。打开 8 连排试管条的盖子，并将 PCR 架面朝上放置在鼓风（杂交）烘箱中，在 45℃下保持约 75 min。

    如果烘箱的热风口偏向 PCR 架一侧，每 15 min 转动架子一次，使得蒸发均匀。必须减少 Re-PCR 产物的体积，以便加入乙醇和醋酸钠沉淀 DNA。

32. 用 150 μl 预冷无水乙醇和 5 μl 3mol/L 醋酸钠（用多道移液器转移这些溶液）沉淀回收 PCR 产物。盖上 8 连排试管并倒转架子数次以混匀。
33. 在－20℃下放置 15 min 沉淀 PCR 架中的样品。
34. 4℃下以 1699 *g* 离心 PCR 架 90 min 来回收沉淀的样品。
35. 使用多道移液器小心除去上清。然后用多道移液器向每个 DNA 团块加上 150 μl 70% 乙醇，盖上试管。
36. 涡旋每个 8 连排试管条直至 DNA 团块变疏松。所有 8 连排试管条涡旋后，将架子在 4℃下以 1699 *g* 离心 45 min。

379

37. 使用多道移液器尽可能多地除去 70% 乙醇。DNA 团块在室温下干燥约 90 min。

    干燥需要的时间取决于团块中残留乙醇的量。不要将团块彻底干燥。

38. 每次仅操作一个 8 连排试管条，使用多道移液器向每个 DNA 团块加入 12 μl 20% DMSO 溶液，并用吸头吹打混匀。盖上试管重悬 DNA 团块，轻弹 PCR 管底部使 DNA 团块变松。

    点样溶液的最终 DNA 浓度应约为 0.8～1.3 μg/μl。

39. 将 DNA 溶液在 4℃下放置过夜（或更长时间）。使用多道移液器混匀溶液，直至沉淀完全重悬。

    DNA 溶液可以在 4℃保存。在点样液中 DNA 已被变性，因此不需更进一步处理。

点样和微阵列的处理

40. 按照第 2 章，方案 2 的介绍进行 DNA 样品点样，然后让其在空气中干燥。

    点样溶液印到玻璃表面之后，由于点样液中不含盐分，样点会变得几乎不可见。对玻片呼气将使得阵列在短时间内可看见。

41. 将点印了阵列的玻片放入 UV Stratalinker 2400 中，阵列向上，暴露于 260 000 μJ UV 照射下，将 DNA 交联到基底上。

阵列应使用固定剂量的能量（260 000 μJ），而不是 Stratalinker 中的其他可用选择，如自动交联或时间。玻片交联过度可能会导致荧光杂交信号减弱。

使用 DAPI 染色玻片可以显示出被点样针沉积后通常保留在点中的全部 DNA，以及在 50% 甲酰胺/2×SSC 杂交中仍然被保留的 DNA。

380 方案 3

BAC 基因组微阵列 CGH

Richard Segraves[*]，**Donna Albertson**[*†‡]，**Daniel Pinkel**[*‡]（[*]综合性癌症中心、[†]癌症研究所、[‡]实验医学系，加州大学，旧金山，加利福尼亚，94143—0808）

在单拷贝增益和丢失的检测中，克隆在 BAC 载体中的基因组 DNA 组成的阵列单元比较小的克隆（如 cDNA）产生更强的信号和更好的性能。用克隆在高容量 BAC 和 PAC（P1-衍生的人工染色体）载体中的基因组插入片段的序列标签（sequence tag），就可以直接获得基因组序列。用随机引物法标记的基因组 DNA 为实验提供了足够的扩增量，以至于 3 ng 的 DNA 就可以用来在 BAC 阵列上测定 CGH。

下面的方案介绍了实验和参考基因组 DNA 样品的制备及荧光标记、样品与 BAC 衍生微阵列杂交的条件，以及随后的图像和数据分析。关于制备 BAC 克隆用于微阵列点样，见方案 2。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

封闭溶液

见步骤 10。

DAPI（4′，6-diamidino-2-phenylindole）染色封片剂

1 μmol/L DAPI	1 μl
** PBS，不含 Ca^{2+} 和 Mg^{2+}	100 μl
甘油	900 μl

EDTA（0.5 mol/L，pH 8.0）

乙醇（无水的），冰冷

据报道，有些批次的乙醇（100%或 200 标准强度）中含有会使微阵列试验中产生很深荧光背景的污染物质，95%的乙醇不存在这个问题。为避免这个问题，可以用 95%的乙醇或者

** 用 0.5mol/L 碳酸盐缓冲液将 PBS 的 pH 调至 8.0。

用已证明不会产生高荧光背景的无水乙醇。Warner-Graham 公司（Cockeysville，Maryland）出售的无水乙醇（USP 级乙醇，64-17-5）效果很好。

甘油（100%）

杂交混合物（pH 7.0）

1 g 硫酸葡聚糖（钠盐，500 000 m. w.）（超纯，Invitrogen）

5 ml 甲酰胺〈!〉

1 ml 20× SSC

将 1 g 粉状硫酸葡聚糖分散在 15 ml 试管的整个内壁上。将试管平放，喷入5 ml甲酰胺。盖上试管并剧烈振摇 30 s。加入 1 ml 20× SSC 并剧烈振摇 30 s。将试管在室温下保存过夜，让使硫酸葡聚糖苷溶解。用水将溶液体积调节到 7 ml，并将 pH 调至 7。

10× 磷酸缓冲液（PBS）（pH 7.4），不含 Ca^{2+} 和 Mg^{2+}

PN 缓冲液

PN 缓冲液由两种组分混合配制：一种二碱式和一种单碱式组分。

- 对于二碱式组分，配制 0.1 mol/L $Na_2HPO_4 \cdot H_2O$
- 对于单碱式组分，配制 0.1 mol/L $NaH_2PO_4 \cdot H_2O$

用单碱式组分将二碱式组分的 pH 调节到 pH 8.0。加入诺纳德 P-40 [Nonidet P-40，NP- 40（乙基苯基聚乙二醇）] 至终浓度为 0.1%。

SDS（20%）〈!〉

Sephadex 溶液

50 g Sephadex G-50

500 ml 10 mmol/L Tris（pH 8）

1 mmol/L EDTA

0.1% SDS

醋酸钠（3 mol/L，pH 5.2）

Tris 碱（1 mol/L，pH 7.6）

洗涤缓冲液（pH 7），预热到 45℃

50% 甲酰胺〈!〉

2× SSC

酶和缓冲液

E. coli DNA 聚合酶 I Klenow 片段（高浓度，40～50 单位/μl）

使用高浓度 Klenow 片段显著改善标记效率。该酶可以作为 BioPrime Labeling Kit（GIBCO）的一部分购买得到，或者从 New England Biolabs 单独购买。

凝胶

琼脂糖凝胶（1%），1×TBE 倒胶

见步骤 6。

核酸和寡核苷酸

Cy5 和 Cy3 标记 dCTP（1 mmol/L 储存液）（PA55021 和 PA53021，Amersham

Biosciences)

10×dNTP-标记溶液：dATP、dGTP、dTTP（每种的浓度为 2 mmol/L）和 dCTP [浓度为 0.5 mmol/L 的 TE（pH 7.6）溶液]

使用超纯 dNTP（如 Pharmacia）。有关 dNTP 储存液的制备及保存的建议，见附录 2。

鲱鱼或鲑鱼精 DNA（10 mg/ml）（Invitrogen）

人 C_0t-1 DNA（0.7 μg/μl 储存液）（Invitrogen）

使用荧光仪测定时，每批新 C_0t-1 DNA 的浓度应为 500 ng/μl 或更高。

BAC 单元组成的微阵列，按照方案 2 中介绍制备。

阵列可点印在多种基底上，包括镀铬玻璃、氨基硅烷包被玻璃和 Corning 的商品化玻片（请见第 2 章中**用于点样的基底**信息栏）。作者一般将阵列点印在铬上，因为这为使用自制的 CCD 成像系统进行 CCD 成像提供了好的环境。对于商品化激光扫描仪，可以将阵列点印在玻璃片上，图像采集无需 DAPI 封片或使用盖玻片。

2.5 ×随机引物溶液（BioPrimer Labeling Kit，18094-011，GIBCO）

125 mmol/L Tris-Cl（pH 6.8）

12.5 mmol/L $MgCl_2$

25 mmol/L β-巯基乙醇〈!〉

750 μg/ml 随机八聚体

这些组分可单独配制，也可以购买完整的缓冲液作为 BioPrime DNA Labeling Kit（Invitrogen）的一部分。

已知浓度的参考 DNA

已知浓度的实验基因组 DNA

使用任何标准方法制备的基因组 DNA 的质量都足以用于阵列 CGH 分析。

酵母tRNA（50 mg）（Invitrogen）

用灭菌水配制 100 μg/μl 溶液。

专用设备

夹子

透明指甲油

盖玻片（No. 1，Corning）

加热槽，预设为 100℃

振荡台（如 VWR 牌 Rocker，Mode 100 ）

橡胶黏合剂（Ross，American Glue Corporation）

Sephadex G-50 离心柱（见步骤 8）（Amersham Biosciences）

硅酮垫圈（按压-封闭，PGC Scientific）

UV Stratalinker 2400

水浴，预设为 45℃

方法

基因组 DNA 随机引物标记

1. 标记两个带旋盖的微量离心管，一个用于基因组 DNA 实验样品，另一个用于基

因组 DNA 参考样品。

每个随机引物标记反应在 50 μl 反应体系中进行，体系中包含 200～300 ng 的基因组 DNA、1×随机引物、40 单位 Klenow 片段、Cy3 或 Cy5 标记的 dCTP（40 μmol/L）和 1×dNTP 混合物。随机引物法产生的 DNA 产物大小一般为 200～1500 bp，其中约 400 bp 片段的浓度最高。见步骤 2-4。

2. 在适当的试管中，加入 200～300 ng 要标记的基因组 DNA 实验和参考样品。加入 20 μl 2.5×随机引物溶液，并用水将每管体积调至 42 μl。
3. 将试管放入加热槽（或热循环仪）中，在 99℃下温育 10 min 使 DNA 变性，然后将试管转移到冰上。

383

DNA 变性之后，如有必要，短暂离心使得内容物集中于试管底部。

4. 每份反应混合物中加入：

dNTP 标记液	5 μl
Cy5 或 Cy3 标记 dCTP	2 μl
Klenow 片段（40 单位/μl）	1 μl

通常，实验基因组 DNA 样品用 Cy5 标记，参考基因组 DNA 样品用 Cy3 标记，Cy-标记 dUTP 可以用来代替 Cy-标记 dCTP，但是，dNTP 标记混合物中核苷酸的浓度应该做相应的调整。

5. 轻敲试管壁混匀组分，并将反应在 37℃下孵育过夜。
6. 每个标记反应取一小部分（5 μl），用 1%琼脂糖凝胶电泳分析。

DNA 电泳条带应当呈现为约 100～500 bp 的扩散带。

7. 使用 Sephadex G-50 柱子，通过离心柱层析除去未掺入的核苷酸。详细方法见 Sambrook and Russell（2001）书中方案 A8.30。

阵列杂交

8. 为了在玻片上生成一个杂交区域，用套在 10 ml 注射器末端的 200 μl 移液器吸头将胶黏剂涂布在阵列周围形成一圈，使用体视显微镜观察该区域（图 6-4）。让胶黏剂在空气中干燥，随后在第一层上面涂上更厚的一层胶黏剂。置于空气中干燥。
9. 混合下列 DNA 样品，制备用于杂交的标记探针：

384

人 C_0t-1 DNA（约 0.7 μg/μl）	100 μl
标记的实验基因组 DNA	50 μl
标记的参照基因组 DNA	50 μl

用标准乙醇沉淀方法回收 DNA 混合物。4℃下以 20 800 *g*（14 000 r/min）离心 30 min 使 DNA 成团块状沉淀，弃去上清，并让 DNA 团块在空气中干燥 5～10 min。

10. 将 DNA 团块溶解于 5 μl 酵母 tRNA 溶液和 10 μl 20 % SDS 中，然后加入 35 μl 杂交混合物。
11. 按照下述为每个阵列制备 50 μl 封闭液：
    a. 用 2.5 倍体积乙醇和 0.1 倍体积 3 mol/L 醋酸钠（pH 5.2）来沉淀 50 μl 鲑鱼精 DNA（10 mg/ml）。混合溶液，在微型离心机中以 7750 *g*（6400 r/min）离心 1～2 min 回收沉淀。

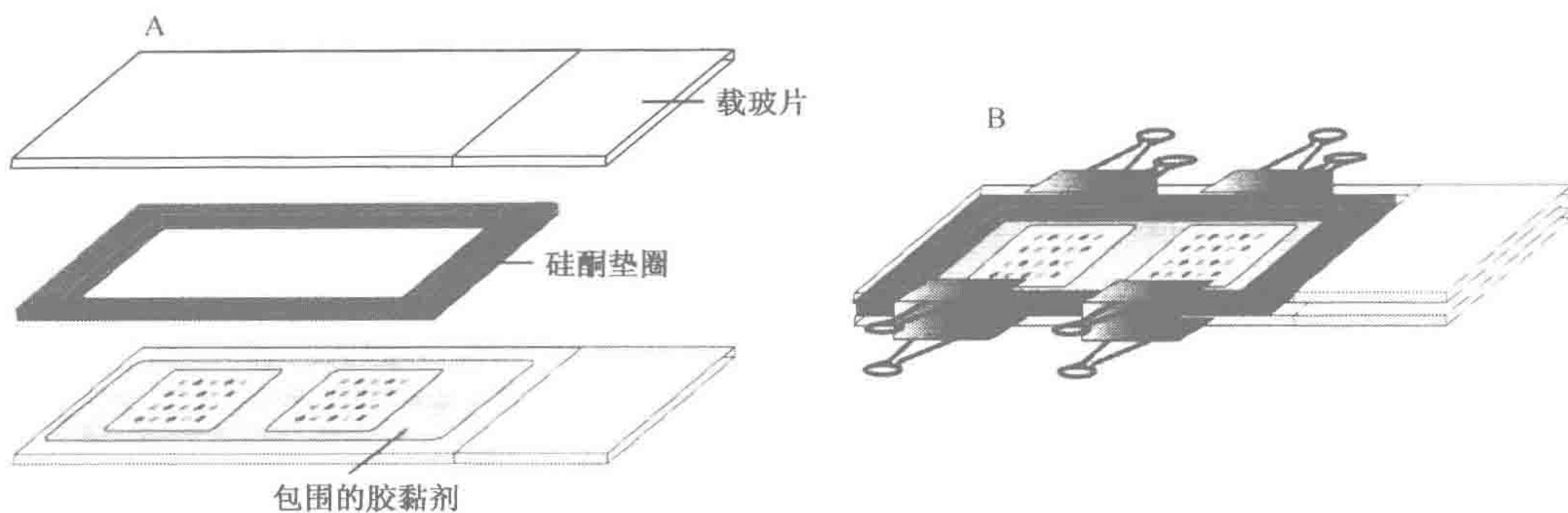


图 6-4　杂交玻片组件概要

杂交液加入包围在阵列周围的胶黏剂圈中，然后将硅酮垫圈放在胶黏剂周围。将一块洁净的显微镜载玻片放在硅酮垫圈上，并将整个载玻片组件如图所示用夹子封紧。A. 放大视图，显示了硅酮垫圈底部阵列周围环绕了胶黏剂的阵列载玻片，和一片用做上盖的载玻片；B. 有夹子的装配好了的杂交室

（经允许，修改自 Snijders et al. 2002）

b. 弃去上清，并让团块在空气中干燥。

c. 将 DNA 团块溶解于 5 μl H_2O 和 10 μl 20% SDS 中，然后加入 35 μl 杂交混合物（pH 7.0）。

12. 将硅酮垫圈放在阵列和胶黏剂圈周围，将封闭液（50 μl）加到杂交区域中，注意使封闭液覆盖整个杂交区域。将一片洁净的显微镜玻璃载玻片盖在垫圈上，以防止蒸发。将阵列在室温下封闭约 30 min。

    见图 6-4，注意本步骤中不需要夹子。

13. 将 DNA 在 73℃下温育 10～15 min 变性，然后在 37℃下继续温育 1 h。

    预复性步骤可以使 DNA 中的高度重复序列复性，这样它们不会再与阵列单元中的相似序列杂交。重复序列封闭失败会导致荧光比率被严重压缩。

14. 从封闭的微阵列上除去玻璃载玻片，小心吸除封闭液，然后将步骤 13 制备的杂交混合物迅速转移到阵列上的杂交区域内。

15. 将一片玻璃载玻片放在垫圈上，并用夹子将载玻片夹紧形成气密的杂交室（图 6-4）。

16. 将玻片“室”放在 37℃培养箱中的振荡台上，缓慢摇动（约 1 r/min），孵育 34～48 h。

17. 停止杂交反应后，从阵列上去除夹子和覆盖的载玻片，用 PN 缓冲液液流洗去阵列上的杂交液。

18. 将带有杂交后阵列的玻片放入装有 50 ml 洗涤缓冲液的 Coplin 染缸中，用水浴预热至 45℃，温育微阵列 15 min。

19. 将微阵列载玻片放入一个新的，装有 50 ml PN 缓冲液的 Coplin 染缸中，然后将染缸在室温下放置 15 min。

20. 用镊子小心除去玻片上的胶黏剂（用 PN 缓冲液使阵列保持湿润），并将玻片转移到新的 PN 缓冲液中，放置 15 min。

21. 从Coplin染缸中取出玻片，并使 PN 缓冲液“滚落”玻片。

重要事项：勿使玻片完全干燥。

22. 用 DAPI 染玻片，以使用自制 CCD 系统对玻片成像。

    a. 将约 200 μl DAPI 染色封片剂加到阵列上。

    b. 用 24×50 No. 1 Corning 盖玻片盖在阵列区域上，从边缘吸去过量的 DAPI 封片剂，并用透明指甲油密封。

    现在阵列可以用于成像。图像可以用商品化的 CCD 或激光扫描系统（如 Axon scanner 4000B，见第 3 章中表 3-9）采集。关于成像更多的讨论，见下面阵列成像及分析信息栏。

    使用商品化激光扫描仪时（见第 3 章，表 3-9），阵列用无菌水洗涤并在台式离心机中离心干燥。

阵列成像及分析

点样在透明基底，如包被的载玻片上的阵列可使用商品化激光扫描仪成像。它们一般是在干燥状态下成像。作者使用点样在镀铬玻片上的阵列，并用自制的 CCD 系统成像，获得了较好的测量效果（Pinkel et al. 1998）。阵列用盖玻片覆盖，封片剂为含染色阵列样点 DNA 的 DAPI 的 90% 甘油/PBS 介质。成像系统可分别获得四种荧光染料的图像，通常，DAPI 用于观察样点，荧光素以及 Cy3 和 Cy5（或光谱与它们等同的），用于同时比较最多 3 种不同的基因组 DNA。使用商业化软件如 Axon Instruments 的 Genepix 来对图像进行定量分析以获得比率，如第 7 章，PART Ⅱ 中介绍的方法，或者用自编程序 UCSF SPOT（Jain et al. 2002）完成。UCSF SPOT 程序使用 DAPI 图像自动寻找所有的点，并建立它们的边界，这样不管样点会出现怎样的不规则形状，都可以让边界适合每个样点。然后，程序会为每个点计算局部背景、样点内每个像素对应的每种荧光基团在背景修正后的强度、荧光比率和诊断参数。作者用每个点的实验信号和参考信号总强度的比率作为最终数据。这种分析基本上是完全自动化的——操作人员的干预只允许确证样点边界是否被恰当地找到和当发现问题时进行小的调整。分析一个约 7500 个样点的阵列在标准的笔记本电脑上只花费不到 1 min 的时间。如果只能得到实验图像和参考图像，比如从激光扫描仪获取的图像就是这种情形，UCSF SPOT 可使用任何一种，或者两者之和以代替 DAP 图像来寻找样点。第二种自编程序利用 UCSF SPOT 的输出，将每个样点和克隆 ID 以及作图信息关联起来，平均重复的点，应用质量控制标准，最后输出可以用于解释的数据。

方案 4

386

基于微阵列的 DNA 拷贝数代表性分析：靶 DNA 制备

Robert Lucito，Michael Wigler（冷泉港实验室，冷泉港，纽约，11724）

与阵列 CGH 一样，代表性微阵列分析可测定两个样品之间 DNA 含量的差异。该方法起源于一种鉴定两个基因组之间序列差异的方法，称为代表性差异分析（representational difference analysis，RDA）（Lisitsyn and Wigler 1993；Lisitsyn et al. 1993，

1995)。

由于人类基因组的复杂性，在从总基因组 DNA 制备的标记靶分子中，单拷贝序列的丰度非常低。当使用与人类基因组同样复杂的靶分子与微阵列杂交进行 DNA 拷贝数的比较测量时，单个探针只能获得低的信噪比，除非探针很大，比如通常的 BAC 或 YAC（yeast artificial chromosome，酵母人工染色体），降低靶分子复杂性的方法之一是优先扩增小的限制性片段以产生基因组代表（Lucito et al. 1998）。在这个方案中（图 6-5），DNA 用限制性内切核酸酶切割，然后与作为 PCR 扩增引物结合位点的接头进行连接。在标准 PCR 中，只有小片段（一般小于 1 kb）被高效扩增，因此扩增的材料是总基因组的一个代表，而不是忠实的拷贝。代表性水平取决于所使用的限制性内切核酸酶（Lucito et al. 1998）。例如，使用识别 6 个碱基序列的酶 *Bgl* Ⅱ 可以得到低复杂性的代表（约为基因组的 2.5%），因此产生小片段的频率相对很少。比较而言，由于

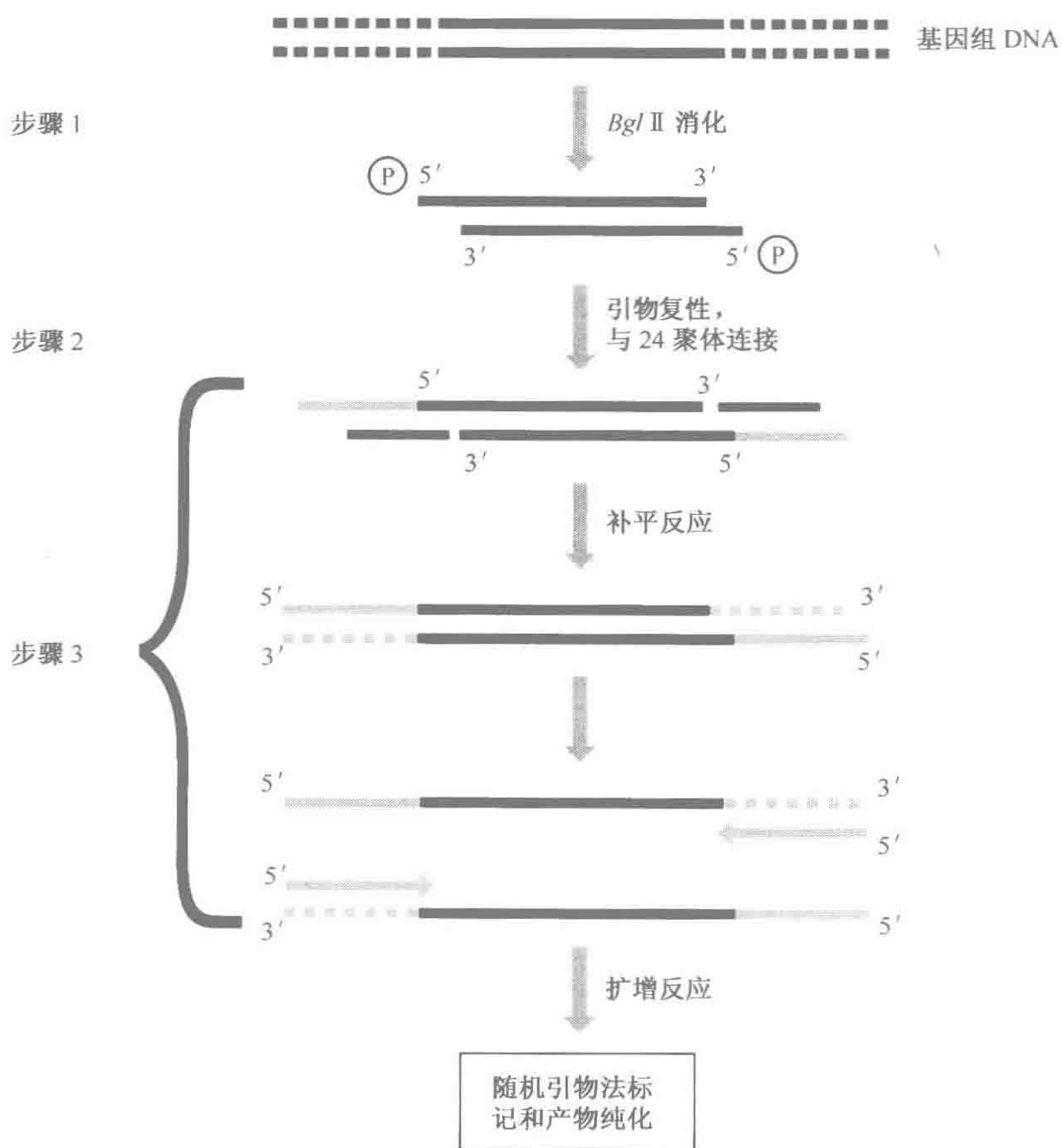


图 6-5　代表性差异分析（RDA）示意图

在步骤 1 中，用限制性酶 *Bgl* Ⅱ 消化基因组 DNA，产生一组带突出 5′端的片段。在步骤 2 中，未被磷酸化的寡核苷酸接头连接到末端。接头由一个 24 聚体和一个部分互补的 12 聚体组成。24 聚体能够连接到 *Bgl* Ⅱ 切割基因组 DNA（步骤 1）产生的磷酸化 5′端上。由于用做引导 24 聚体的 12 聚体在 5′端没有磷酸基团，因此不能连接到基因组 DNA 片段的 3′端。在随后的 PCR 中（步骤 3），12 聚体被置换掉，凹陷的 3′端被补齐，并且基因组 DNA 片段被扩增

识别4个碱基序列的酶 *Dpn* Ⅱ，由于频繁切割而产生很多小片段，因此得到高复杂性的代表（约70%）。在低复杂性代表中，每个标记成分的贡献比例要比在含有整个基因组片段的混合物中高。因此，使用低复杂性的标记代表集能够改善信噪比。从两种DNA来源（如肿瘤和对应的正常组织）制备的标记代表集可以分别用不同的荧光染料区分标记，一般是Cy3和Cy5。标记的靶分子与微阵列进行竞争性杂交，来检测两个样品DNA拷贝数的差异（Lucito et al. 2000）。

用代表性微阵列方法分析时，微阵列的组成是重要的考虑因素。因为代表性靶分子来自于基因组的随机片段，对于编码和基因间隔序列没有选择。因此，含有cDNA探针的微阵列不是非常适合这种方法。作者使用了含有两种来源探针的微阵列进行他们的研究：

一个选择是使用随机分布在整个基因组中的基因组小片段（小于1.0 kb），或者是，在某些情况下，从感兴趣的特定区域克隆（Lucito et al. 2000）。另外一类探针，目前受到作者的偏爱，是长链寡核苷酸（约70 bp）。本方案中讨论了这些探针的相对优点。 **387**

材料 **388**

注意：标记有〈!〉的材料的正确处理方法参见附录3。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录2。

使用前将储存液稀释到合适的浓度。

10×扩增缓冲液（N808-0191，Applied Biosystems）。

作为替代，10×扩增缓冲液也可按照以下配方配制：

500 mmol/L KCl

100 mmol/L Tris-HCl（pH 8.3）

将缓冲液按照小等份分装，在−20℃下保存。

乙醇(无水，70%)

据报道，有些批次的乙醇（100%或200标准强度）中含有会使微阵列试验中产生很深荧光背景的污染物质，95%的乙醇不存在这个问题。为避免这个问题，可以用95%的乙醇或者用已证明不会产生高荧光背景的无水乙醇。Warner-Graham公司（Cockeysville，Maryland）出售的无水乙醇（USP级乙醇，64-17-5）效果很好。

异丙醇〈!〉

$MgCl_2$（25 mmol/L）

酚：氯仿（*V*/*V*）〈!〉

SDS（10 %）〈!〉

醋酸钠（pH 5.3）

SSC

3×SSPE

SSC和SSPE可自行配制或从一些公司，如Invitrogen或Ambion购买不含RNase的

产品。

TE（pH 8.0）

酶和酶缓冲液

E. coli DNA 聚合酶Ⅰ Klenow 片段（5 μl）（M0210S，New England Biolabs；或 18012 〈!〉 021，GIBCO）

注意 Klenow 由 Megaprime DNA Labeling System 提供。该方案将需要额外的酶（步骤 11），可以向别的供应商订购。

Megaprime DNA Labeling System，dCTP（RPN1607，Amersham Biosciences）

该随机引物法标记试剂盒与标记的 dCTP 同时使用。

限制性内切核酸酶（New England Biolabs）

选择适合于消化基因组 DNA 的内切酶（本方案中使用 BglⅡ作为示例），参见步骤 1。

Taq 聚合酶（5 单位/μl）（Perkin Elmer 或 Applied Biosystems）

T4 DNA 连接酶（4 单位/μl）（New England Biolabs）

凝胶

琼脂糖（2%）凝胶，自配的或预制胶（BioWhittaker）

389 核酸和寡核苷酸

接头

以下列出的引物对可以从任何信誉良好的寡核苷酸合成公司订购。

200 nmol 的 12 聚体足够进行连接反应（步骤 3），除非计划制备很多组代表集。需要 1 μmol的 24 聚体，因为它同时被用作 PCR 扩增引物（步骤 6）和连接反应（步骤 3）。将引物以 62 pmol/μl 的浓度重悬在水中，小份分装后在－20℃下保存。

以下是三对可供选用的接头：

RBgl24	agcactctccagcctctcaccgca
RBgl12	gatctgcggtga
NBgl24	aggcaactgtgctatccgagggaa
NBgl12	gatcttccctcg
IBgl24	tcagcatcgagactgaacgcagca
IBgl12	gatctgctgcgt

以上任何一对接头都可用于本方案中的连接反应。因为在不同批次的引物中观察到有细微差异，所有相关联的实验中应该使用同一来源的引物。对于构建寡核苷酸阵列，只需使用一对接头。对于从代表物制备片段（RDA）或从基因组克隆，如 YAC 或 BAC 中制备片段，为减少污染或交叉杂交的机会，推荐更换使用的接头。

重要事项：这些寡核苷酸是用于和由 *Bgl*Ⅱ限制性内切核酸酶消化产生的片段的末端进行连接的。可对这些接头序列做适当修改以适合任何可以生成突出 5′端的限制性内切核酸酶。

载体 tRNA（5～10 μg）（Sigma）

C_ot-1 DNA

可选的，见步骤 12。

Cy3 和 Cy5 标记的 dCTP（Amersham Biosciences）

dNTP 溶液，含全部四种 dNTP，每种的浓度均为 10 mmol/L（Sigma）

基因组 DNA

基因组 DNA 可以从样品或用做参照的正常样品中分离得到（可从 CLONETECH 或 Sigma 获取）。

专用设备

离心蒸发器（如 SpeedVac）

加热槽，设定为 55℃

杂交室（TeleChem International）

Microduster 罐装压缩气体除尘器（VWR）

Microcon YM-30（Amicon）

带热盖的热循环仪（MJ Reserch 或 Applied Biosystems）

方法

扩增子代表的制备（图 6-6）

1. 用限制性内切核酸酶消化基因组 DNA。

   本方案中以 *Bgl*Ⅱ为例，但是，也可以使用其他的酶。酶的选择会影响基因组的代表程度（Lucito et al. 1998）。当消化少于 500 ng 的 DNA 时，反应要在体积至少为 50 μl，并且存在 20 ng/μl tRNA 的情形下进行。为减少代表物中的干扰，要确保消化进行完全。 **390**

   让消化反应在 37℃下温育至少 8 h 或者过夜。正常 DNA 及肿瘤 DNA 来自于同一个体是非常重要的，否则位于 *Bgl*Ⅱ位点的多态性会导致个体之间代表性的差异。用于消化的基因组 DNA 的量取决于可获得的 DNA 数量情况。通常，如果使用肿瘤标本，数量可以少于 25 ng。

2. 用酚：氯仿抽提 1 次，纯化消化的片段，加入载体（5～10 μg tRNA），2 倍体积无水乙醇，以及 1/10 体积的 3 mol/L 醋酸钠（pH 5.3）。涡旋样品以混匀，并在冰上放置 15 min。

3. 在微型离心机中，以最大速度（14 000 *g*）将试管离心 15 min，弃去上清，用 70% 乙醇洗涤沉淀团块。弃去乙醇，在 SpeedVac 中干燥团块至阴干。 **391**

   仔细操作，不要丢失团块。

4. 将 DNA 重悬于 10 μl TE 中，并将消解的 DNA 和接头连接。建立如下连接反应：

纯化的 DNA 消化物（步骤 3）	10 μl
24 聚体接头（62 pmol/μl）	7.5 μl
12 聚体接头（62 pmol/μl）	7.5 μl
连接酶缓冲液	3 μl

5. 为促进复性，缓慢冷却反应。取出试管，在冰上放置数分钟。向反应液中加入 2 μl T4 DNA 连接酶（4 单位/μl），在 16℃下反应过夜。

   一种简便的缓慢降温方法是：把反应置于 55℃的加热槽中，然后把加热槽放在冷的房间里直到加热槽的温度降到 16℃。

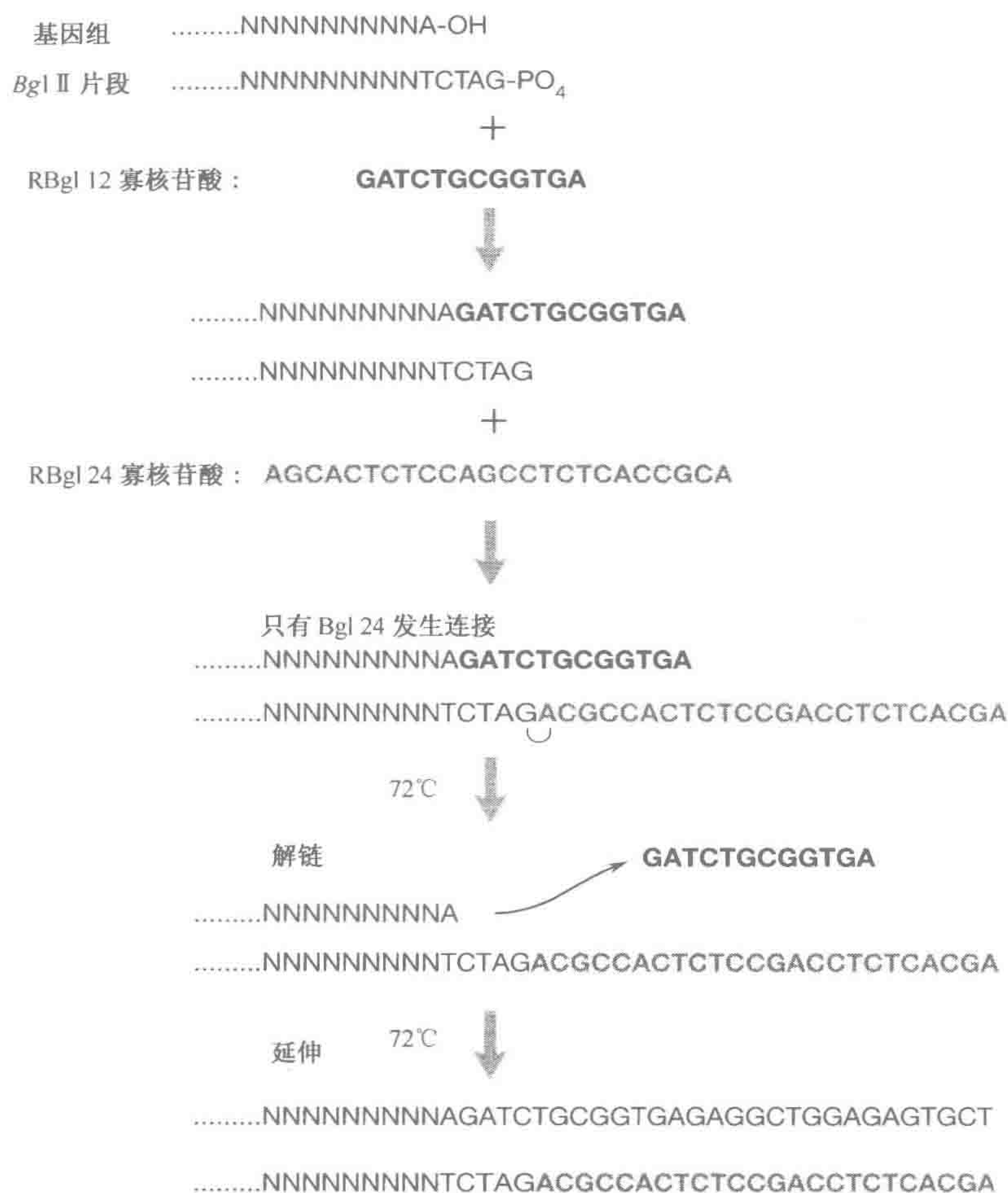


图 6-6 RDA 连接和扩增中接头的应用

较短的寡核苷酸（本例中的 RBgl12）与 *Bgl*Ⅱ位点互补，引导较长的寡核苷酸（RBgl24）加入。只有 RBgl24 通过基因组片段上的 PO_4 基团发生连接

6. 使用适合于起始材料数量的条件扩增连接反应。以下一系列的步骤适合于扩增 100 ng 到 1 μg 模板 DNA（较少量 DNA 的扩增条件在步骤 c 下方的信息栏中给出）：

   a. 将连接反应的 1/10 分别加入两个 PCR 管中。向每管加入以下试剂：

H_2O（终体积 400 μl）	304.2 μl
扩增缓冲液	40 μl
$MgCl_2$（25 mmol/L）	32 μl
dNTP 溶液	12.8μl
接头	8 μl
Taq 聚合酶（5 单位/μl）	3 μl

$MgCl_2$ 的浓度对代表物的样式至关重要。$MgCl_2$ 浓度越高，代表中扩增的某一特定长度

以上的片段越长。作者发现在本方案中 2 mmol/L 的镁浓度最适于 Perkin Elmer *Taq* 酶。凝胶上的代表集的样式或大小范围，在相同浓度 $MgCl_2$ 下随不同的 *Taq* 聚合酶而改变。如果使用 Perkin Elmer *Taq* 之外的其他聚合酶，最好将不同浓度 $MgCl_2$ 条件下得到的结果和使用在含有 2 mmol/L $MgCl_2$ 缓冲液中的 Perkin Elmer *Taq* 聚合酶取得的结果进行比较。只有少数热循环仪，如 ABI 480，能进行 400 μl 体积的 PCR 反应。如果没有合适的仪器，反应可以分成数管进行。

b. 把试管放入热循环仪中，在 72℃下温育 5 min，使得凹陷的 3′端补齐。

c. 按照以下的温度和时间扩增 DNA 模板。

循环次数	变性	聚合
扩增循环（25）	95℃ 1 min	72℃ 3 min
最终循环		72℃ 10 min

一次反应的平均产量为 15～20 μg 扩增的 DNA。

392

少量模板 DNA 的扩增

对于如从少量细胞分离得到的少量 DNA 的扩增，需要增加一个额外的步骤以提高扩增产物的产量，

- **扩增 25～100 ng DNA**：按照上面步骤 a、b 和 c 介绍的体系及扩增条件，只进行 20 个循环的扩增（步骤 c）。此时，将每个反应分成两管，加入适当体积的扩增混合物（步骤 a 准备的 PCR 试剂），将每份反应的体积调至 400 μl，再进行另外的 5 个 PCR 循环。
- **扩增从激光俘获显微切割得到的 1000～5000 个细胞**：使用从步骤 5 开始的整个连接反应，按照上面步骤 a、b 和 c 介绍的体系及扩增条件，只进行 15 个循环的扩增。此时，将每个反应分成两管，加入适当体积的扩增混合物（步骤 a 准备的 PCR 试剂），将每份反应的体积调至 400 μl，再进行另外的 10 个 PCR 循环。

7. 用酚：氯仿抽提纯化 PCR 产物。加入 1/10 体积的醋酸钠和 1 体积异丙醇，将反应在冰上放置 15 min，然后在微型离心机中以最大速度（14 000 *g*）离心 15 min。弃去上清，用 70% 乙醇洗涤团块。弃去上清，并在 SpeedVac 中干燥团块至阴干。
8. 将代表集重悬在原体积 1/10 的 TE 中。取出 1 μl，用 2%琼脂糖凝胶电泳分析产物。

   代表集中片段大小应为 200～1200 bp。代表应与浓度已知的鲱鱼精 DNA 标准品进行比较，来估计代表集的浓度。当从人 DNA 制备 *Bgl* Ⅱ 代表时，在代表集中有两个明显的条带出现，大小约为 600 bp 和 800 bp。它们是人基因组中重复序列的扩增产物。如果代表集扩散模糊，或者这些条带没有出现，最好重新制备代表集。

标记和纯化

9. 使用商品化试剂盒，用随机引物法标记代表集。尽管其他公司的随机引物法试剂盒也能满足需要，还是建议使用 Amersham Biosciences 的试剂盒。按照试剂盒中介绍的反应条件，并做下列修改：

a. 在一个离心管内混合下列试剂：

代表集	8～12 μl
随机引物（由试剂盒提供）	10 μl
随机九聚物（由试剂盒提供）	5 μl

b. 将离心管在 95℃下放置 3 min。随后短时离心进行混合，并将试剂集中在管底。

393 c. 往试管中加入下列试剂：

Cy5 或 Cy3 dCTP（根据不同实验）	8～10 μl
标记混合物（由随机引物试剂盒提供）	20 μl
Klenow 片段（由随机引物试剂盒提供）	4 μl

d. 将试管在 37℃下温育 1 h。

e. 加入 4 μl Klenow 片段（来自其他供应商），并将反应在 37℃再次温育 1 h。

10. 用 Microcon YM-30 柱纯化标记的代表集。

a. 将两管在柱上合并，一个代表集用 Cy5 标记，另一个代表集用 Cy3 标记，并向柱中再加入 150 μl TE。

b. 室温下，在微型离心机中，将柱子以 12 500 *g* 离心 10 min。

11. 除去柱中流出液体，向残留物（标记的代表集）中再加入 400 μl TE，室温下，在微型离心机中将柱子以 12 500 *g* 离心 10 min。除去柱中流出液体并重复该步骤 3 次。

如果靶分子要与含有基因组片段、而非寡核苷酸探针的微阵列杂交，要加入 C_0t-1 DNA 以封闭非特异性杂交。如果需要，C_0t-1 DNA 应该在步骤 11 中最后一次洗涤之前加入。如果在后面加入，标记样品的最终体积将会增加到不能接受的水平。

12. 残留样品的体积应当约为 5～7 μl。倒转柱子并将其放入一支新的离心管中。室温下将试管以 12 500 *g* 离心 2 min 收集标记的样品。

13. 向收集的样品中，加入 3× SSPE 使体积为 35 μl，并加入 SDS 至浓度为 0.2 %（0.7 μl 10% SDS）。

此处，标记的靶分子可以立即与微阵列杂交（见方案 5）。

394

方案 5

基于微阵列的 DNA 拷贝数代表性分析：玻片制备及杂交

Robert Lucito，Michael Wigler（冷泉港实验室，冷泉港，纽约，11724）

使用源自低复杂性代表物的寡核苷酸探针进行点样制备的微阵列与基因组 DNA 的代表杂交，来检测和鉴定导致拷贝数改变的差异（扩增和缺失）。下面介绍了将已经标记和扩增的代表集与微阵列进行杂交的方案。该方案建立和进行数据分析的特别注意事项在方案最后的疑难问题解答中给出。

作者使用了两种来源的探针构建 RDA 微阵列：

- **克隆的小于 1.0 kb 的人 DNA 限制性片段**。这些片段来自含有人目的 DNA 区域的 BAC、P1 和 YAC 克隆切除的小的 *Bgl2* Ⅱ 片段、RDA 片段文库或者人 DNA 的 *Bgl2* Ⅱ 片段的随机克隆（Lucito et al. 2000）。
- **与分布在整个人类基因组中特定序列匹配的长链寡核苷酸**。可以参考人 DNA 序列（http://genome.ucsc.edu/）来得到长链寡核苷酸。有很多原因使得作者更加偏爱长链寡核苷酸，而不是限制性片段。寡核苷酸在受控制的条件下合成，因此，产量要比那些 PCR 扩增片段更加趋于均一。这样，寡核苷酸阵列上样点之间的变异通常要比 PCR 片段阵列的小。另外，还能够避免各种与序列相关的问题，如限制靶分子杂交的重复序列、单核苷酸延伸（mononucleotide stretch）和发夹结构。对染色体序列对应的寡核苷酸知识的了解促进了与文献中介绍的各种肿瘤中增益或丢失区域的相互比较。最后，通过计算机（in silico）设计，产生多个在基因组中相邻的寡核苷酸，从而能够对来自一簇探针的实验结果进行平均。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

乙醇(无水，冰冷的)

据报道，有些批次的乙醇（100%或 200 标准强度）中含有可以使微阵列试验产生很深荧光背景的污染物质，95%的乙醇不存在这个问题。为避免这个问题，可以用 95%的乙醇或者使用已鉴定证明不会产生高荧光背景的无水乙醇。Warner-Graham 公司（Cockeysville, Maryland）出售的无水乙醇（USP 级乙醇，64-17-5）效果很好。

0.2% SDS〈!〉

0.2% SDS/1×SSC

0.2×SSC

0.05×SSC

核酸和寡核苷酸

DNA 微阵列，按照第 2 章，方案 2 介绍的方案点样在载玻片上

用于微阵列的玻片是硅烷化处理的，购自 Corning [Corning MicroarrayTechnology (CMT)（http://www.corning.com/CMT/）：γ-氨基丙基硅烷（GAPS)]。其他玻片表面需要使用不同的方法进行预处理。阵列通常用浓度为 35 pmol/L 的寡核苷酸点样。

基因组 DNA，按照第 2 章，方案 4 介绍的方法进行标记

专用设备

盖玻片

许多种盖玻片可应用于阵列杂交，但作者建议使用 lifter slips（来自 Erie Scientific Co.）或

Hybrislips（来自 Grace Biolabs，HS40）。

杂交室（TeleChem International）

Microduster（VWR）

玻片盒（60-6441-00，PGC Scientific 或者类似的产品）

附加试剂和设备

步骤 1～3 需要第 2 章，方案 2 中列出的试剂和设备。

步骤 5～11 需要第 3 章，方案 20 中列出的试剂和设备。

396

方法

玻片的点样后处理

关于 RDA 微阵列点样后处理的详见方法，见第 2 章，方案 2。同样的方案基本上可以用在此处，并作如下修改：

1. 将微阵列在湿盒内重新水化 2～3 min。从盒中取出玻片，在 60 mJ 的 Stratlinker 紫外交联仪中进行交联固定。
2. 把玻片在 100～140℃的热台上加热约 5 s 使载玻片快速干燥，然后将它们在空气中冷却。
3. 用 0.2% SDS 洗涤玻片 10～15 s，然后转移入蒸馏水中洗涤 10～15 s。把玻片在沸腾的蒸馏水中浸没 3 min，使 DNA 片段变性。将玻片在冰冷的乙醇中浸没数秒钟，然后离心甩干。

   如果阵列由长链寡核苷酸构成，就不需要在沸水中变性。
4. 将处理好的载玻片存放在闭光的盒子中，如无软木衬垫的塑料玻片盒中。

杂交

关于杂交的详细情况，见第 3 章，方案 20。同样的方案基本上可以用在此处，并作如下修改：

5. 把玻片放入杂交室中，在每张玻片上盖上盖玻片。
6. 为使靶分子变性，将标记好的样品在沸水中处理 5 min。沸水处理之后不要把样品放置在冰上，以防杂交液中的 SDS 沉淀。

   在微量离心管上使用盖锁（如 LidLock，PGC Scientific），以防靶分子变性时盖子被弹开。样品蒸发会改变 SSPE 和 SDS 的浓度，影响杂交条件。
7. 短时离心，收集凝集在管盖上的材料。

   如果已经加入了 C_0t-1 DNA 来封闭重复序列，就将探针在加到阵列上之前将探针在 37℃下温育 30 min。如果使用寡核苷酸，在设计时应该已经避免了重复序列。因此就不需要封闭处理了。
8. 向位于 Telechem 杂交室两条短边上的小池内加入 10～15 μl 水。沿盖玻片周围滴加探针，利用毛细作用使液体在盖玻片下扩展开来。

   避免在向微阵列上加入靶分子和盖玻片时产生气泡，以及防止微阵列在温育过程中变干的方法，见第 3 章，方案 20。

9. 盖上杂交室的盖子，并将杂交室放入杂交炉中，在 65℃下温育 12～18 h。 397

洗涤

10. 按照以下的温度和时间洗涤玻片：

0.2% SDS/1×SSC	55℃	直至盖玻片从载玻片上漂离
0.2% SDS/1×SSC	55℃	1 min
0.2% SDS	20℃	30 s
0.05% SDS	20℃	10 s

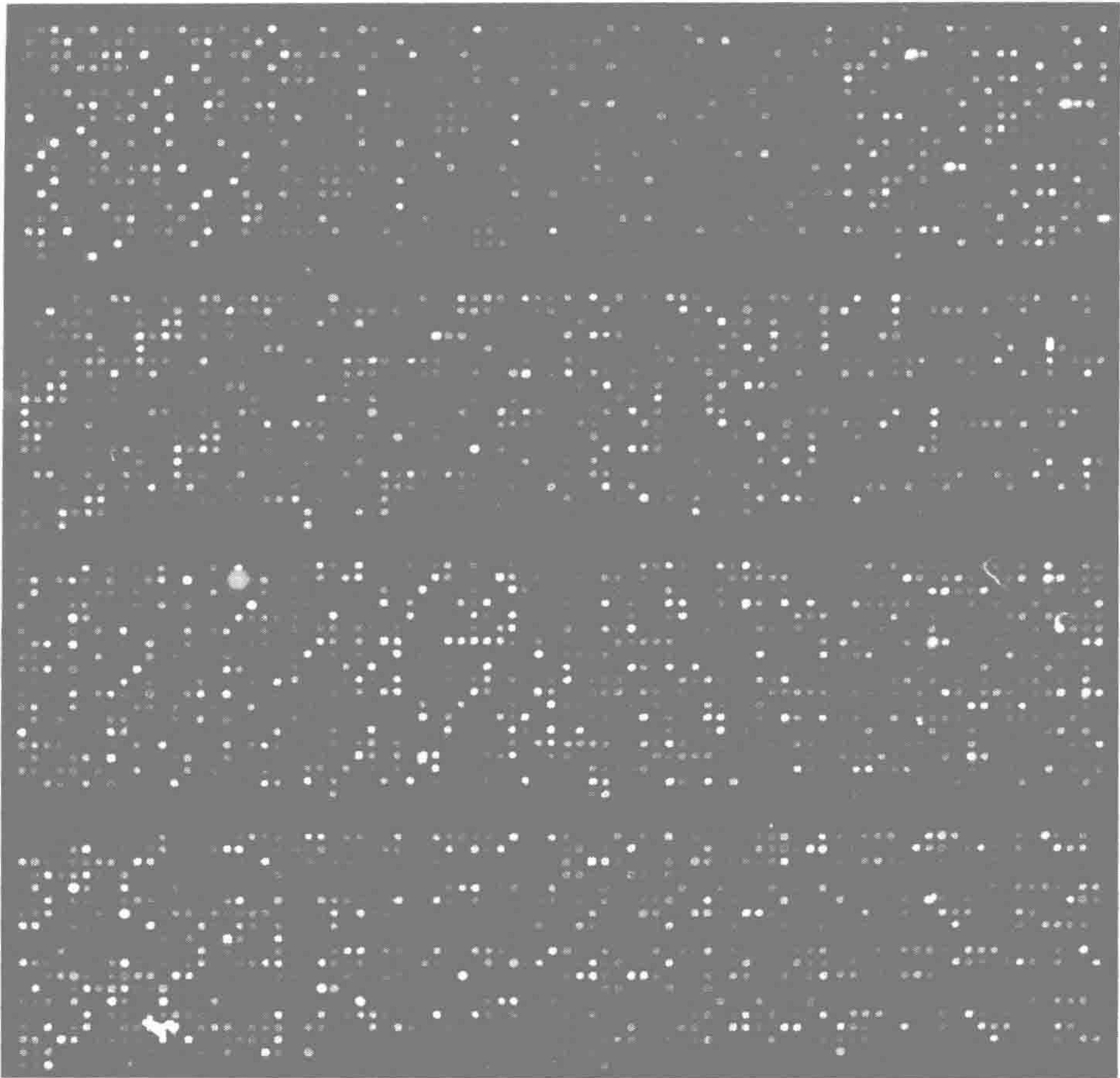


图 6-7 原发性乳腺癌标本基因拷贝数的微阵列分析

从乳腺癌活检标本中制备细胞核，根据基于 DNA 倍性的荧光激活细胞分选进行分离。本例比较中，二倍体核视为正常，非整倍体核视为肿瘤的。从两种类型的核中分离基因组 DNA，制备代表集，进行差异标记，与寡核苷酸阵列杂交，然后清洗和分析。代表集的标记方式可以使拷贝数增益的探针（在肿瘤中被扩增）显示为红色，而拷贝数丢失的探针显示为绿色（见彩图）

玻片可以逐片清洗，也可以将它们放在玻片架上成组清洗。在第一种洗涤液中，小心移去盖玻片。需要轻微摇晃以促使盖玻片从载玻片上脱离，如果盖玻片没有立即滑落，就将玻片在洗涤液中停留更长的时间。盖玻片脱落困难可能是由于杂交过程中液体蒸干而造成的。 398

在后续实验中，可以通过向纯化的标记样品中加入2～3 μl 3×SSPE 来增加杂交体积从而补救这个问题。

警告：在洗涤的任何阶段，玻片不能干燥。不能使用过量的 SDS，在扫描之前必须将其完全除去，以避免背景过高。

11. 将玻片在台式离心机中低速离心甩干。

玻片可以放在玻片架上，也可以放在管底塞有数块无绒擦布的 50 ml 离心管中。

图 6-7 显示了典型的杂交结果。

分　析

有许多软件包可以计算阵列中样点或者特征单元的强度（见第 7 章，PART Ⅳ）。作者目前使用的是 Axon 扫描仪中包含的软件。强度一得到测定，软件就计算 Cy5/Cy3 的比率。数据被保存下来并且可用其他软件作进一步分析。进一步的分析用来标准化数据，使得中值比率为 1.0（基因组中半数以上的基因应该没有拷贝数变化）。用于分析的软件包（如 Spotfire 和 SPlus）应该能够将拷贝数的变化定位到特定的染色体序列上。Spotfire 软件的优点是方便用户的，可以在分析过程中随时灵活地改变阈值标准，而且该软件有多种聚类算法和多种察看数据的方法。Spotfire 一个有趣的功能是信息是相互关联的，通过选择图表中的一个点，使用者可在多个其他实验中追踪该特性单元的行为。SPlus 使用较难，但是它极其灵活，而且可以对数据进行复杂的统计学分析。SPlus 还有一种脚本语言，能够编写脚本来执行各种功能，包括导人数据、数据的标准化及和合并多个实验。

399

疑难问题解答

代表集和标记

- **制备代表集。**使用相同的方法制备阵列中的不同代表集是很重要的。作者比较了在不同的试管中，不同的时间，和用不同的消化反应制备的代表集。要进行比较的代表集应该使用相同的接头，因为在使用不同的接头制备代表集时，有时会观察到大的差别。通常，当两种要比较的样品在同一天制备，并且在同一个 PCR 仪中使用相同的接头时，会得到最佳的结果。
- **阵列上信号强度低。**如果过高地估计了代表集的浓度，或者使用的代表集的数量少于推荐值，会导致信号强度偏低。弱信号也可能是由代表集标记效率低引起的，标记效率低可能是由于 Klenow 片段未能有效地延伸，或者是核酸或引物的质量差。可以用更换标记反应中的全部组分的方法来解决这一问题。

杂交反应（见第 3 章中的疑难问题解答指南）

- **扫描的阵列中出现过亮和过暗的区域。**如果玻片清洗不正确或者清洗后没有通过离心充分干燥，常常会出现这些情况。一个可能的解决方法是增加清洗的次数和（或）增加清洗后的离心时间。
- **扫描的阵列中出现污斑。**这个问题通常是由玻片的清洗不够引起的。如果该问题持续存在，在准备时需要增加额外的玻片清洗时间。污斑也可能是由清洗过程中 SDS 去除不充分而引起的。增加最后一次清洗的时间可以缓解这个问题。
- **阵列上有明亮的边缘。**如果杂交混合物的体积太小，使得盖玻片下方液体流动不畅，会出现这个问题。增加杂交混合物的体积，直至少量的液体残留在盖玻片的外面。杂交混合物体积太大会使盖玻片漂走！

- **阵列的边缘或角落出现较暗的区域。**该问题通常是由盖玻片弯曲引起的。阵列越大，盖玻片越大，弯曲越厉害，从而导致阵列将杂交液从玻片上挤走，使得杂交不均匀。解决的方法是制作小一些的阵列，或者使用较厚的盖玻片。
- **整个阵列中出现污渍。**这个问题可能是由玻片表面包被不均匀引起的，并且如果在点样结束时，使用很高品质 PMT 的扫描仪进行扫描，也能够检测出该问题。

阵列的质量差

- **阵列上的样点很小。**当玻片没有被正确地重新水化时，会产生这个问题。重新水化时适当地增加时间。
- **样点大小有变化。**这种样点大小的变化同样由一到两个原因引起。如果 PCR 扩增后使片段沉淀，并且用于沉淀的盐没有被充分的除去，残存的盐分留在探针中。在重新水化时，含盐较多的样点会比含盐少的样点膨大更多。有两个解决方法：①洗板时更加仔细以除去残留的盐；②通过柱层析纯化片段。作者实验室的成员通常使用 Millipore Multiscreen 过滤板从引物、核苷酸和蛋白质等试剂中纯化片段，并且，纯化过程使用可编程的自动工作站。实践证明该系统工作良好，唯一的不足是会损失约 30%的材料。样点大小发生变化也可能是由于干燥之后未能很好地重悬阵列材料。重悬的体积小以及任何的移液误差都会导致盐浓度的差异和上述的问题。与样品制备有关的点样问题一般不会在寡核苷酸集出现。
- **样点中的 DNA 浓度低。**无论这个问题仅限于特定的样点，还是整个多孔板都受到影响，很显然的可能原因是扩增不足或者后续纯化中的损失。这个问题可以在点样之前通过凝胶分析解决。

PART Ⅱ 突变检测及 SNP 基因分型

400

导言

两项最近的进展唤起了哺乳动物基因组分析与作图方面的意义深远的变革：丰富的单核苷酸多态性（single-nucleotide polymorphisms，SNP）的大型信息数据库的发展，包括这些 SNP 的等位基因在目标群体中以大致相同的频率出现以及构建包含数千个寡核苷酸探针的、高度有序的、可寻址的阵列技术的发展。应用这些阵列，从大量的 SNP 中构建等位基因型（allelotype）的问题就变成了简单的笛卡儿几何学问题。

SNP 是在一个物种的个体之间最为常见的信息丰富的遗传改变。在公共数据库中，SNP 的数目已超过 160 万，这相当于在人类基因组中，大约每 2000 kb 中有一个 SNP 的频率。除了密度高以外，与其他类型的遗传标记相比，SNP 还具有很多优点，例如，与短的二核苷酸和三核苷酸重复相比，SNP 具有在进化上的稳定性，以及它们可以适用于各种与 DNA 长度无关的评分方法。SNP 等位基因的识别方法是通过等位基因特异性杂交或者等位基因特异性引物延伸来实现，而不是测量特定 DNA 片段大小的差异（Kwok 2000；Shi 2001）。将一种等位基因识别方法与一种检测方案，如荧光强度、荧光偏振、微阵列结合或质谱相结合，产生了多种可行的高通量基因分型方案。更为常用方法的总结，见表 6-2。从原理上讲，这些方法可以用来对大量基因组中数千个 SNP 进行评分。在这个水平上的工作在单基因疾病连锁研究、特定基因型与常见疾病的关联研究（多基因疾病）以及遗传药理学研究都是十分必要的。

由于需要对大量个体中的数千个 SNP 进行评分（Chakravarti 1998；Collins et al. 1998；Wang et al. 1998），高密度微阵列非常适合于 SNP 基因分型。非毗邻的遗传学

表 6-2 基因分型的常用方法

方 法	等位基因识别	检 测	多重性	参考文献
限制性片段长度多态性（RFLP）	消化	片段大小，凝胶电泳	否	Kiko et al.（1979）
等位基因特异性扩增	延伸	存在与否，凝胶电泳	是	Fortina et al.（1992）
*Taq*Man	杂交	荧光，溶液	否	Livak et al.（1995）
分子信标	杂交	荧光，溶液	否	Tyagi et al.（1998）
入侵检查（the invader assays）	切割	荧光，溶液	否	Ryan et al.（1998）
寡核苷酸连接检查（OLA）	连接	荧光，凝胶电泳	是	Eggerding（1995）
荧光共振能量转移（SBE-FRET）	延伸	荧光，溶液	否	Chen and Kwok（1999）
荧光偏振（SBE-FP）	延伸	荧光，溶液	否	Chen et al.（1999）
多重长度加尾引物单碱基延伸（LM-SBE）	延伸	荧光，凝胶电泳	是	Lindblad-Toh et al.（2000b）
DASH	杂交	荧光，溶液	否	Prince et al.（2001）
焦磷酸测序	延伸	荧光，溶液	否	Ronaghi et al.（1999）
质谱（MALDI-TOF）	延伸	片段大小，飞行时间	是	Haff and Smirnov（1997）
SBE 与流式细胞术结合	延伸	荧光，流式细胞术	是	Chen et al.（2000）
微阵列上等位基因特异性引物延伸	延伸	荧光，阵列	是	Pastinen et al.（2000）
阵列上的单碱基延伸（SBE-TAG）	延伸	荧光，阵列	是	Hirchhorn et al.（1998）
人类 SNP 芯片	杂交	荧光，阵列	是	Sapolsky et al.（1999）
Affymetrix 标签阵列	延伸	荧光，阵列	是	Fan et al.（2000）

多态性可以按照高度平行化的方式进行分析，这样就可以对可能在某个 SNP 位点处出现的各种等位基因进行识别。在基于微阵列的基因分型中，等位基因的识别可以直接在阵列上（等位基因特异性杂交）或是在预杂交过程中（单碱基延伸）进行。在这两种情况下，由于有效的检测通常需要远远超过 fmol 量级的靶分子 DNA，而且含有目的 SNP 的序列要用基因组 DNA 作为模板并通过多重 PCR 进行扩增。因此，阵列可以采用预先合成的寡核苷酸点样或者是通过光刻法在阵列上直接合成探针来制备。

包含有多组覆盖了目的基因组区域的叠瓦式探针（tiled probe）的微阵列，也可以被用来检测一段特定序列中已知的或可预测的改变和突变。扩增靶基因组 DNA 片段，并与阵列进行杂交。杂交模式（hybridization pattern）显示了靶样品中存在的特定序列。

本部分包含四个方案：

- **方案 6：SBE-标签阵列用于 SNP 基因分型**。通过 PCR 扩增 SNP 并与位点特异性引物复性。然后以荧光标记的双脱氧核苷酸三磷酸盐（ddNTP）作为底物进行单碱基延伸反应（SBE）。扩增的 SNP 等位基因特定的掺入了其中一种标记 ddNTP，然后通过延伸引物中的标签序列与点样微阵列中基因组探针之间的杂交对其进行检测。
- **方案 7：Affymetrix 标签阵列用于 SBE SNP 基因分型**。与方案 6 类似，该方案介绍了基于应用通用高密度寡核苷酸 Affymetrix 阵列进行的 SNP 基因分型。GeneFlex Tag

阵列由通用的20目寡核苷酸标签探针组成，在原理上，它与标记寡核苷酸的杂交可以用来检测任何SNP。

- **方案8：高密度微阵列用于序列变异检测**。对感兴趣的靶区域中每个碱基进行检测或再测序的策略，可以通过合成含有一系列固定长度的叠瓦式寡核苷酸的阵列来实现。与靶序列互补的所有可能寡核苷酸都出现在阵列上。其中每条探针依次由一组至少四个变异体来代表，每个变异体在设计的查询位置上（例如，接近探针中心位置的一个碱基）都含有一个不同的碱基（A、G、C、T）。

  目的基因组中扩增的、荧光标记的靶分子与阵列中和它完全互补的寡核苷酸之间的杂交最为牢固。方案中介绍了应用该技术检测人类*p*53基因的突变。 **402**

- **方案9：SNP阵列用于杂合性缺失分析**。该方案中通过与Affymetrix GeneChip HuSNP进行等位基因特异性杂交来识别在人类肿瘤中由于缺失而形成纯合的DNA区域。以生殖细胞（正常的）DNA和从肿瘤中抽提的DNA作为模板，用5′加尾的位点特异性寡核苷酸作为引物进行多重PCR扩增。每个被扩增的位点都与Affymetrix GeneChip中显示的一个SNP相对应。将第二次生物素化修饰的扩增产物合并在一起，与SNP阵列杂交，并用链亲和素-藻红蛋白耦联物进行染色。通过对比生殖细胞DNA与肿瘤DNA的杂交模式，检测杂合性缺失位点。

方案6 **403**

SBE-标签阵列用于SNP基因分型

Kerstin Lindblad-Toh[*]，Yin-Mei Lim[*]，Joel N. Hirschhorn[*†‡]，Pamela Sklar[*†§]（[*]怀特黑德研究所/MIT基因组研究中心，剑桥市，马塞诸塞州 02115；[†]哈佛医学院，波士顿，马塞诸塞州 02115；[‡]遗传及内分泌科，儿童医院，波士顿，马塞诸塞州 02115；[§]精神病科，麻省总医院，波士顿，马塞诸塞州，02114）

在基于微阵列的基因分型中，等位基因识别可以直接在阵列中进行，或者是在阵列杂交之前的预备步骤中进行。Affymetrix HuSNP GeneChip包含将近1500个可同时进行基因分型的人类SNP位点，是在阵列上直接进行等位基因识别的一个例子（Sapolsky et al. 2000）。玻璃载玻片上的单碱基延伸标签阵列（single-base extension-tag arrays on glass slides，SBE-TAGS方法）是等位基因识别在杂交前发生的一个例子（见下述）（Hirschhorn et al. 2000）。

直接在微阵列上进行基因分型有些不够灵活，因为等位基因的数量和种类是由微阵列的组成决定的。但是，通过用双重功能的SBE-TAG引物在杂交前进行等位基因识别，人们可以得到一种“通用的”微阵列，而且还保留了在SNP检测中的灵活性。这些引物携带有：提供识别等位基因的独特的位点特异性序列；与微阵列中特定寡核苷酸“标签”互补的序列。这样，一组标签识别序列可与不同组的位点特异性序列相互组合，从而达到等位基因识别的灵活性。这里给出的SBE-TAGS方案是基于Hirschhorn等（2000）所介绍的方法。在Affymetrix GeneFlex芯片（Fan et al. 2000）中使用了一种与之相似的标签策略（见方案7）。

标签阵列的制备方法可以是点样预合成的寡核苷酸（标签寡核苷酸），如同本方案中所介绍的，或者是通过光刻法直接在阵列表面合成标签（如 GeneFlex 芯片，方案 7）。光刻合成法产生均匀的样点，但是这种方法成本高，并且在标准的实验室中不可行。与此相比，如果标签引物点样均匀并且浓度相等，在玻片上点样寡核苷酸不仅成本低廉而且容易进行。点样材料的量可以通过向每种标签寡核苷酸点样混合物中加入阳性对照寡核苷酸来监测。阳性对照引物包含一个特异性的序列，通过与互补的荧光标记的寡核苷酸杂交来检测。在本方案中，阳性对照用 FAM 标记，其荧光输出在扫描仪的一个特别通道中读出。应用阳性对照可以对样点的值进行标准化，以补偿点样效率的变化。

在常规的 SBE 中（Syvanen 2001），用位于 SNP 位点两侧的引物来 PCR 扩增含有 SNP 的位点。这些引物从合适的数据库中的 DNA 序列中选取。例如，db-SNP 数据库（www. ncbi. nlm. nih. gov/SNP/）提供了关于人类基因组中 SNP 的位置和它们两侧序列的信息。SNP 位点经 PCR 扩增之后，进行产物纯化以除去未掺入的 dNTP 和未使用的引物。随后，一个位点特异性 Tag-SBE 引物，其终止于距离目的 SNP 的 5′端一个碱基的位置，与 PCR 产物复性。在进行的 SBE 反应中，使用 DNA 聚合酶催化模板特异性的荧光标记 ddNTP 的掺入。与荧光标记终止物循环测序一样，每种核苷酸标记有一种不同的荧光染料。由于在延伸反应中没有 dNTP，只有与模板链上 SNP 互补的双脱氧核苷酸才会掺入。通常，SBE 反应产物通过荧光测序设备或者多孔板阅读器来检测。

图 6-8 是 SBE-TAGS 方案的示意图。SBE-TAGS 过程中的第一个步骤与通常的 SBE 方案相似，但是标记产物的识别方法不同。每条 SBE-TAGS 引物设计有一个尾部，与点样在微阵列中的一条标签反向互补（reverse complement）。当 SBE-TAGS 产物被合并，并与微阵列杂交时，来自每个位点的 SBE 产物将会与微阵列中包含相应标签的特异样点发生结合。随后检查每个样点的荧光颜色并确定基因型（图 6-8）。PCR 反应可以是单重的也可以是多重的，单重 PCR 比较简单且不易出错，多重 PCR 节省 DNA 模板的开支并可以减少 PCR 反应的次数。主要方案中介绍了单重 PCR SBE 反应，在替代方案中介绍了多重 PCR 方法。

在设计标签寡核苷酸以及掺入到 tag-SBE 引物中的互补序列时要非常小心。标签寡核苷酸不仅在彼此之间要具有独特性，在进行基因分型的生物体基因组中也要是独特的。表 6-3 列出了 166 条为下述方案研究并经过测试的标签。这些标签选自独特的噬菌体 λ 的序列，具有一致的长度、GC 含量、解链温度（T_m）以及低的自我相似性，不包含两个以上连续相同的碱基（Hirschhorn et al. 2000）。同时，也通过向 BLAST 检索提交标签序列，与已知的序列进行比对来选择相互之间，以及与人类基因相似性低的标签。标签寡核苷酸设计的详细情况见本导言最后 **SBE-TAGS 的设计与使用**信息栏。生成标签的一个替代方法是考虑具有特定长度的全部可能的 *n*-聚体（*n*-mer），例如，20 目，然后选择那些序列独特，而且自我互补性低的标签。在 Affymetrix 标签阵列（方案 7）中使用了这样的一组标签序列（Fan et al. 2000）。

SBE-TAGS 方法的一个重要特性是通过将许多 SNP 位点合并于一次 SBE 和（或）PCR 反应中，而具有高度的多重性（multiplexing）。当进行多重反应时，非常关键的是要确保合并物中的位点不会包含可能产生交叉杂交的位点，从而导致非特异性扩增的同源片段。根据引物/位点之间的相似程度，将它们分配到不同的合并物中，可以避免

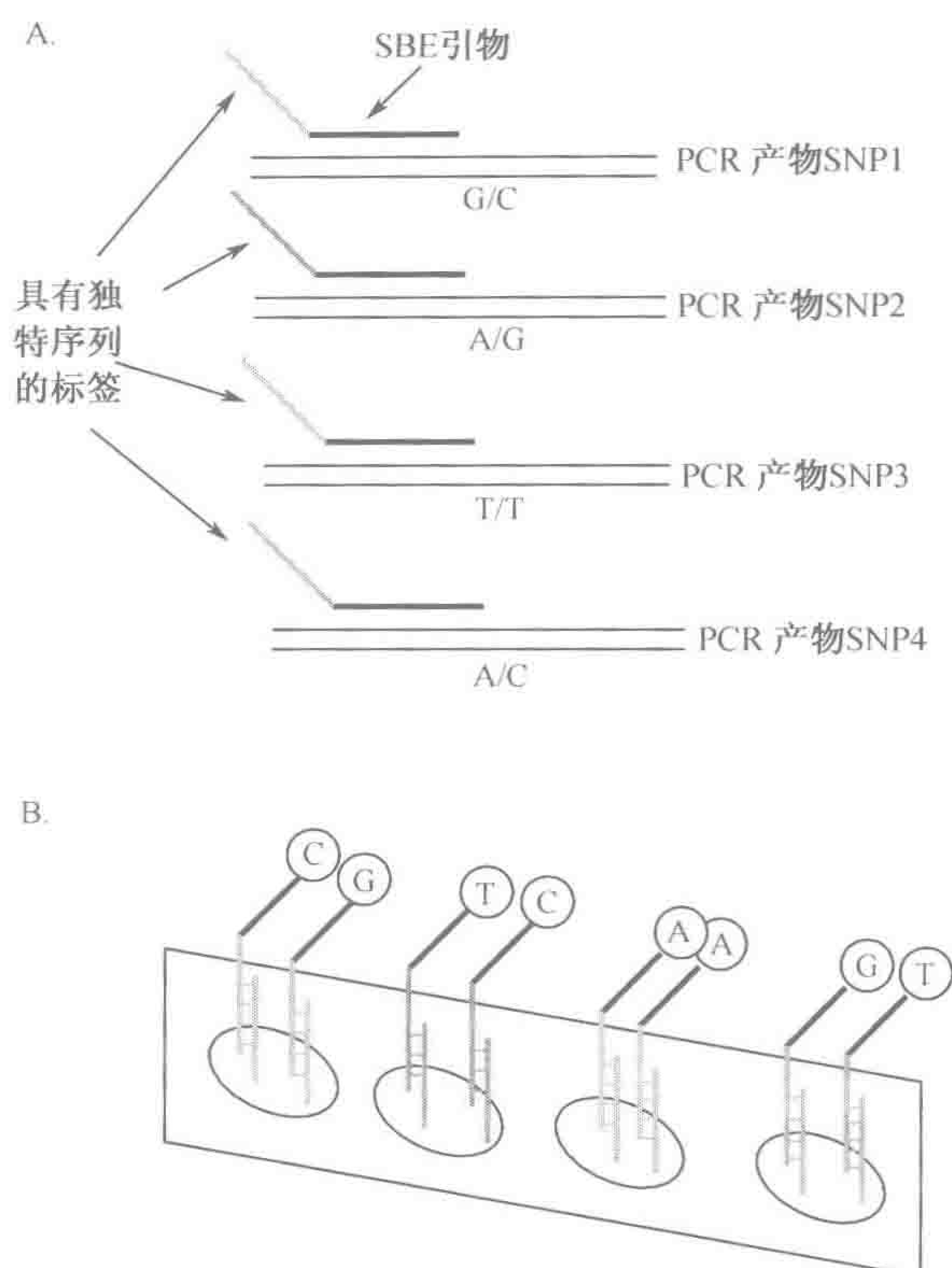


图 6-8　SBE-TAGS 在基因分型中应用示意图

A. 在 SBE 中，一段包含了通用序列标签和紧随其后的位点特异性序列的混合式引物（hybrid primer）与 SNP 附近的位置杂交，并在荧光双脱氧核苷酸存在的情况之下进行延伸。四条延伸产物中的每条都携带有独特的荧光标记，用不同颜色的尾部表示。B. 在溶液中进行多重 SBE 反应，每条 SBE 引物由不同的独特序列标签所标记。与使用序列标签的反向互补序列点样在玻璃载玻片上制备的通用标签阵列杂交后，分析多重反应

［经允许复制，来自 Hirschhorn et al. 2000（© National Academy of Science，U. S. A.）］

这个问题的产生。

设计 Tag-SBE 引物时的一个关键目标是避免标签识别与位点特异性序列之间的自我互补。如果标签-SBE 引物 3′端的一些碱基与寡核苷酸的另一部分区域之间互补，引物就会产生自身折叠，形成发夹结构。在这种情况下，将由引物中的下一个碱基，而不是 SNP 位点来决定延伸产物，从而得到的基因分型结果在很大程度上是错误的。

在下面的方案中，合成的 Tag-SBE 引物与表 6-3 中的标签序列反向互补，SBE 引物序列紧接其后。关于 Tag-SBE 引物设计更多的详细情况在本导言最后的信息栏中给出。

总结

405

本方案修改自 Hirschhorn 等（2000）。主要步骤及其目的总结如下：

- **标签寡核苷酸和 SBE-TAGS 引物的设计**。一个需要仔细计划的关键步骤。参考表 6-3 之后 **SBE-TAGS 的设计与使用**信息栏。

表 6-3　用于 SBE 的噬菌体λ标签

标签	标签序列：用于在载玻片上点样	反向互补序列：用于结合到 SBE 引物中	标签	标签序列：用于在载玻片上点样	反向互补序列：用于结合到 SBE 引物中
LT1	GCCTCCACGCACGTTGTGATATGTA	TACATATCACAACGTGCGTGGAGGC	LT42	CATCGTCAACGACGTTCTCATGGTT	AACCATGAGAACGTCGTTGACGATG
LT2	GCAGATCAATTAATACGATACCTGCG	CGCAGGTATCGTATTAATTGATCTGC	LT43	CCATTCCAGACATGCTCGTTGAAGC	GCTTCAACGAGCATGTCTGGAATGG
LT3	GGTTCTGTTCTTCGTTGACATGAGG	CCTCATGTCAACGAAGAACAGAACC	LT44	GCCATCGCTGGACTATCGAAGAGTG	CACTCTTCGATAGTCCAGCGATGGC
LT4	TTAGTCTCCGACGGCAGGCTTCAAT	ATTGAAGCCTGCCGTCGGAGACTAA	LT45	ATCTCGTTCCGTATCGCGTCGAACT	AGTTCGACGCGATACGGAACGAGAT
LT5	CTGTGACAGAGCCAACACGCAGTCT	AGACTGCGTGTTGGCTCTGTCACAG	LT46	GAACGCAATATTCACAAGCAATGCG	CGCATTGCTTGTGAATATTGCGTTC
LT6	CCTGGTGGTTGACTGATCACCATAA	TTATGGTGATCAGTCAACCACCAGG	LT47	AAGAGACCGCGACTTACCATGTATC	GATACATGGTAAGTCGCGGTCTCTT
LT7	GCATGTATAGAACATAAGGTGTCTC	GAGACACCTTATGTTCTATACATGC	LT48	GCAGAACTGATGAGCGATCCGAATA	TATTCGGATCGCTCATCAGTTCTGC
LT8	GCTAGATGAAGAGCAAGCGCATGGA	TCCATGCGCTTGCTCTTCATCTAGC	LT49	AACCTTCAACTACACGGCTCACCTG	CAGGTGAGCCGTGTAGTTGAAGGTT
LT9	TACAACCGACAGATGTATGTAAGGC	GCCTTACATACATCTGTCGGTTGTA	LT50	GTGTGGCGAGACAGCGACGAAGTAT	ATACTTCGTCGCTGTCTCGCCACAC
LT10	TTCAATCTGGTCTGACCTCCTTGTG	CACAAGGAGGTCAGACCAGATTGAA	LT51	CCTGTATCAGGACATGGTACGAGCTA	TAGCTCGTACCATGTCCTGATACAGG
LT11	ACACGATGTGAATATTATCTGTGGC	GCCACAGATAATATTCACATCGTGT	LT52	AAGGAGATTATGTACCGAGGAAGAA	TTCTTCCTCGGTACATAATCTCCTT
LT12	TTGAAGTTCGCAGAATCGTATGTGT	ACACATACGATTCTGCGAACTTCAA	LT53	AATACCGGAACATCTCGGTAACTGC	GCAGTTACCGAGATGTTCCGGTATT
LT13	AACGTCTGTTGAGCACATCCTGTAA	TTACAGGATGTGCTCAACAGACGTT	LT54	TTGAGAAGGAAGATATCCTCGCATGG	CCATGCGAGGATATCTTCCTTCTCAA
LT14	GGCAACTCATGCAATTATTGTGAGC	GCTCACAATAATTGCATGAGTTGCC	LT55	GGAATTACACCACGTGGATTGGCATC	GATGCCAATCCACGTGGTGTAATTCC
LT15	CCAGAAGTATATTAATGAGCAGTGCAG	CTGCACTGCTCATTAATATACTTCTGG	LT56	GGATATACCGCTCACCGTATTGCAG	CTGCAATACGGTGAGCGGTATATCC
LT16	AAGCAGTCTGTCAGTCAGTGCGTGAA	TTCACGCACTGACTGACAGACTGCTT	LT57	TCATGGATACAGGTTGTGAACATCCAA	TTGGATGTTCACAACCTGTATCCATGA
LT17	AATGATGCTCTGCGTGATGATGTTG	CAACATCATCACGCAGAGCATCATT	LT58	TTGCGGAGCTATTAGAGCTTATACA	TGTATAAGCTCTAATAGCTCCGCAA
LT18	AATACACGAAGGAGTTAGCTGATGC	GCATCAGCTAACTCCTTCGTGTATT	LT59	TTGTGGTAATAGGCCAGTCAACCAG	CTGGTTGACTGGCCTATTACCACAA
LT19	GCTGTTAATCATTACCGTGATAACGCC	GGCGTTATCACGGTAATGATTAACAGC	LT60	GGCCTGCTTAGGTGACGTCTCTCGT	ACGAGAGACGTCACCTAAGCAGGCC
LT20	GCGGAACGGTCAGAGAGATTGATGT	ACATCAATCTCTCTGACCGTTCCGC	LT61	AACCAATCGTAGTAACCATTCAGGA	TCCTGAATGGTTACTACGATTGGTT
LT21	GTTATGGTCAGTTCGAGCATAAGGC	GCCTTATGCTCGAACTGACCATAAC	LT62	AAGATGACCATCTACATTACTGAGC	GCTCAGTAATGTAGATGGTCATCTT
LT22	TTACCTATGATTGATCGTGGTGATATCCG	CGGATATCACCACGATCAATCATAGGTAA	LT63	CGCAGATGAGCTTGTCCATATGACTG	CAGTCATATGGACAAGCTCATCTGCG
LT23	GCTGTGGCATTGCAGCAGATTAAGG	CCTTAATCTGCTGCAATGCCACAGC	LT64	AACGTCAGCGAGCTGGTTGATATGG	CCATATCAACCAGCTCGCTGACGTT
LT24	TGACGTCATTGTAGGCGGAGAGCTA	TAGCTCTCCGCCTACAATGACGTCA	LT65	GAGCCTGGCTAACCGTGACCAGAAC	GTTCTGGTCACGGTTAGCCAGGCTC
LT25	TCAATAATCAACGTAAGGCGTTCCT	AGGAACGCCTTACGTTGATTATTGA	LT66	AACATCGCCGCACAGATGGTTAACTT	AAGTTAACCATCTGTGCGGCGATGTT
LT26	TTATCGGCTACATCGGTACTGACTC	GAGTCAGTACCGATGTAGCCGATAA	LT67	AGAATTCTGGCGAATCCTCTGACCA	TGGTCAGAGGATTCGCCAGAATTCT
LT27	CCATTATCGCCTGGTTCATTCGTGA	TCACGAATGAACCAGGCGATAATGG	LT68	AACGGTGCAATGAAGCCAAGTTAGA	TCTAACTTGGCTTCATTGCACCGTT
LT28	GGCGTACCTTCGCGGCAGATATAAT	ATTATATCTGCCGCGAAGGTACGCC	LT69	AATATGCTGCTTGAGGCTTATTCGG	CCGAATAAGCCTCAAGCAGCATATT
LT29	AACTGAGCCGTAGCCACTGTCTGTCC	GGACAGACAGTGGCTACGGCTCAGTT	LT70	AATACTCAACTTCGGCAGAGGTAAC	GTTACCTCTGCCGAAGTTGAGTATT
LT30	GTTGTGCTGAATTAAGCGAATACCG	CGGTATTCGCTTAATTCAGCACAAC	LT71	AATCCAGATGGAGTTCTGAGGTCAT	ATGACCTCAGAACTCCATCTGGATT
LT31	TTATATCTGCACAACAGGTAAGAGC	GCTCTTACCTGTTGTGCAGATATAA	LT72	AATCGCAATGCTTGGAACTGAGAAG	CTTCTCAGTTCCAAGCATTGCGATT
LT32	AAGAGGCGGCGCTTACTACCGATTC	GAATCGGTAGTAAGCGCCGCCTCTT	LT73	TCTACGAATCGAGAGTGCGTTGCTT	AAGCAACGCACTCTCGATTCGTAGA
LT33	CGGTCACACGTTAGCAGCATGATTG	CAATCATGCTGCTAACGTGTGACCG	LT74	ATCTAACACCGTGCGTGTTGACTAT	ATAGTCAACACGCACGGTGTTAGAT
LT34	AAGAAGAGTCAATCGCAGACAACAT	ATGTTGTCTGCGATTGACTCTTCTT	LT75	GCGTCCTCAAGCTGCTCTTGTGTTA	TAACACAAGAGCAGCTTGAGGACGC
LT35	CATATCGCGCTGTGACGATGCTAAT	ATTAGCATCGTCACAGCGCGATATG	LT76	CACTTCTAAGTGACGGCTGCATACT	AGTATGCAGCCGTCACTTAGAAGTG
LT36	AAGGTACTGCAAGTGCTCGCAACAT	ATGTTGCGAGCACTTGCAGTACCTT	LT77	TCTTCTGAACCAGACTCTTGTCATT	AATGACAAGAGTCTGGTTCAGAAGA
LT37	CTGCATTCTGCGGTAAGCACGAACT	AGTTCGTGCTTACCGCAGAATGCAG	LT78	AATGATTCGTCATCTGCGAGGCTGT	ACAGCCTCGCAGATGACGAATCATT
LT38	AACCTAACATTGATTCAGGTACAGG	CCTGTACCTGAATCAATGTTAGGTT	LT79	CCTCATCAGTGGCTCTATCTGAACG	CGTTCAGATAGAGCCACTGATGAGG
LT39	CAGTTGATCATCAGCAGGTAATCTGG	CCAGATTACCTGCTGATGATCAACTG	LT80	ATGGTTCCTGCATATGATGACAATG	CATTGTCATCATATGCAGGAACCAT
LT40	ATATGTTATCTGCCACGCCGATTAT	ATAATCGGCGTGGCAGATAACATAT	LT81	GGTATCATGTAGCCGCTTATGCTGGA	TCCAGCATAAGCGGCTACATGATACC
LT41	AACTGGATACGATTGGATTCGACAA	TTGTCGAATCCAATCGTATCCAGTT	LT82	GCATATATGAATGAACGATGCAGAG	CTCTGCATCGTTCATTCATATATGC
			LT83	TTGGTAGGTGAGAGATCTGAATTGC	GCAATTCAGATCTCTCACCTACCAA

续表

标签	标签序列：用于在载玻片上点样	反向互补序列：用于结合到 SBE 引物中	标签	标签序列：用于在载玻片上点样	反向互补序列：用于结合到 SBE 引物中
LT84	TTATGTCAACACCGCCAGAGATAAT	ATTATCTCTGGCGGTGTTGACATAA	LT125	AAGTATTCGTTCACTTCCGATAAGC	GCTTATCGGAAGTGAACGAATACTT
LT85	GTGCGTCCTGCTGATGTGCTCAGTAT	ATACTGAGCACATCAGCAGGACGCAC	LT126	CTATGTGCCGGAGCGGACATTACAA	TTGTAATGTCCGCTCCGGCACATAG
LT86	TTGTGCATCCATCTGGATTCTCCTG	CAGGAGAATCCAGATGGATGCACAA	LT127	TCGTTGCTGGAAGCCTGGAAGAAGTA	TACTTCTTCCAGGCTTCCAGCAACGA
LT87	AATGGATCCACTCGTTATTCTCGGA	TCCGAGAATAACGAGTGGATCCATT	LT128	GGAATATCAGAAGTGGAACGGCACA	TGTGCCGTTCCACTTCTGATATTCC
LT88	TTAACCACATCAGGCTCGGTGGTTCT	AGAACCACCGAGCCTGATGTGGTTAA	LT129	TTATAATCTGCTGGCCGGAACTAAT	ATTAGTTCCGGCCAGCAGATTATAA
LT89	AATTGCTATAAGCAGAGCATGTTGCG	CGCAACATGCTCTGCTTATAGCAATT	LT130	GGTGCACATGCGCATACAGTTGGTA	TACCAACTGTATGCGCATGTGCACC
LT90	CACTCAGGCGGCCTTGATAGTCATA	TATGACTATCAAGGCCGCCTGAGTG	LT131	CTCACAGTCTGAGCGGTTCAACAGG	CCTGTTGAACCGCTCAGACTGTGAG
LT91	ACGCCTCGAGTGAAGCGTTATTGGT	ACCAATAACGCTTCACTCGAGGCGT	LT132	CAGCGGTCGTGCTGTATTGTCTCAG	CTGAGACAATACAGCACGACCGCTG
LT92	CAATGGCATATTGCATGGTGTGCTC	GAGCACACCATGCAATATGCCATTG	LT133	CTACCACCATGACTAACGCGCTTGC	GCAAGCGCGTTAGTCATGGTGGTAG
LT93	GCTCTGAGCCTCAAGACGATCCTGAAT	ATTCAGGATCGTCTTGAGGCTCAGAGC	LT134	GTCACCTGACGCACTGAATACGCTG	CAGCGTATTCAGTGCGTCAGGTGAC
LT94	CATGTCGTCTGCCAGTTCTGCCTCT	AGAGGCAGAACTGGCAGACGACATG	LT135	GTACTGGCCGCGATTGCAGATGTTA	TAACATCTGCAATCGCGGCCAGTAC
LT95	TACCTCTGAATCAATATCAACCTGG	CCAGGTTGATATTGATTCAGAGGTA	LT136	AAGATATAGCTTCAGCTGTCGCGCTT	AAGCGCGACAGCTGAAGCTATATCTT
LT96	TTCATCGTTAACCGGAGTGATGTCG	CGACATCACTCCGGTTAACGATGAA	LT137	GGCGTTACAGCATGGATGTGGAGTA	TACTCCACATCCATGCTGTAACGCC
LT97	GGAACAGAGCGGCAATAAGTCGTCA	TGACGACTTATTGCCGCTCTGTTCC	LT138	CAGTGGCGTGGAGTGCAGGTATACAG	CTGTATACCTGCACTCCACGCCACTG
LT98	GGTACGCTGCAGGATAATGTCCGGT	ACCGGACATTATCCTGCAGCGTACC	LT139	GGTGGTCCGGCAGTACAATGGATTA	TAATCCATTGTACTGCCGGACCACC
LT99	CTGTTCGACAGCTCTCACATCGATCC	GGATCGATGTGAGAGCTGTCGAACAG	LT140	CACTATGAGAGCCTGCGTGGACGTT	AACGTCCACGCAGGCTCTCATAGTG
LT100	ACATAATGCAGGCCTTCACGCTTCA	TGAAGCGTGAAGGCCTGCATTATGT	LT141	GGATTAACGTGAAGTACCGTTATGA	TCATAACGGTACTTCACGTTAATCC
LT101	TTCGCGGTGCTTCTTCAGTACGCTA	TAGCGTACTGAAGAAGCACCGCGAA	LT142	GGAAGTAAGCGTACTGTCAGCGGCA	TGCCGCTGACAGTACGCTTACTTCC
LT102	TTACTGATAAGTTCCAGATCCTCCT	AGGAGGATCTGGAACTTATCAGTAA	LT143	AACCAGCACGCGTTATCTTGGTACG	CGTACCAAGATAACGCGTGCTGGTT
LT103	TTCGCTCACTTCGAACCTCTCTGTT	AACAGAGAGGTTCGAAGTGAGCGAA	LT144	AACGCTCTGGTCGTCATACACTGAA	TTCAGTGTATGACGACCAGAGCGTT
LT104	GTCTTGTTCGATTAACTGCCGCAGAA	TTCTGCGGCAGTTAATCGAACAAGAC	LT145	GGAATACCAACATATCCGGTGTCACG	CGTGACACCGGATATGTTGGTATTCC
LT105	CCGTAATAGCGATGCGTAATGATGT	ACATCATTACGCATCGCTATTACGG	LT146	GATTGAAGGTCCGGTGGATGGCTTA	TAAGCCATCCACCGGACCTTCAATC
LT106	GCCAGCGTCAGACATCATATGCAGATAC	GTATCTGCATATGATGTCTGACGCTGGC	LT147	GAAGCGAAGGACAACCTGAAGTCCA	TGGACTTCAGGTTGTCCTTCGCTTC
LT107	TTGAGTACGGTCATCATCTGACACT	AGTGTCAGATGATGACCGTACTCAA	LT148	TTCTCTCGGTGCCAGTATGGTGCTC	GAGCACCATACTGGCACCGAGAGAA
LT108	TGTATGCCGACTCTATATCTATACC	GGTATAGATATAGAGTCGGCATACA	LT149	GCGCAGTTACATGAGACTCTGCCTGA	TCAGGCAGAGTCTCATGTAACTGCGC
LT109	CCATATCATCCAGTGGTCGTAGCAG	CTGCTACGACCACTGGATGATATGG	LT150	GAGCTGCTGCACCATATTCCTGAAC	GTTCAGGAATATGGTGCAGCAGCTC
LT110	CAATTCGAATATTGGTTACGTCTGC	GCAGACGTAACCAATATTCGAATTG	LT151	ATCGTGAGGATGACTGGTGGCGTAA	TTACGCCACCAGTCATCCTCACGAT
LT111	ATAACTCAATGTTGGCCTGTATAGC	GCTATACAGGCCAACATTGAGTTAT	LT152	GCTCTGAAGGCGGTGTATGACATGG	CCATGTCATACACCGCCTTCAGAGC
LT112	ATTATCACTGTTGATTCTCGCTGTC	GACAGCGAGAATCAACAGTGATAAT	LT153	TGAATAACGTCATGTCAGAGCAGAA	TTCTGCTCTGACATGACGTTATTCA
LT113	TTCGTAGCGATCAAGCCATGAATGT	ACATTCATGGCTTGATCGCTACGAA	LT154	AGATTGCGTCCATCAGCCAGAGTGT	ACACTCTGGCTGATGGACGCAATCT
LT114	CCTCGTCTATGTATCCATTGAGCATT	AATGCTCAATGGATACATAGACGAGG	LT155	CGAGTCACTCAGCGCACTGGTTAAG	CTTAACCAGTGCGCTGAGTGACTCG
LT115	TTCCTTCATTGATATTCCGAGAGCA	TGCTCTCGGAATATCAATGAAGGAA	LT156	CAGGAGGAGTCCGGTATGGCTGAAC	GTTCAGCCATACCGGACTCCTCCTG
LT116	CAAGTACTAATAAGCCGATAGATAGCC	GGCTATCTATCGGCTTATTAGTACTTG	LT157	AACATACTGAACTCCGTGCAGCCGT	ACGGCTGCACGGAGTTCAGTATGTT
LT117	TTAACACACGTGCGAACTGTCCATGA	TCATGGACAGTTCGCACGTGTGTTAA	LT158	GGATGATGGCGGAAGTTGTCATCTC	GAGATGACAACTTCCGCCATCATCC
LT118	GCTGGCAGCATTCTTGAGTCCAATA	TATTGGACTCAAGAATGCTGCCAGC	LT159	AGAACTTCCTGCCGGACAACACGAT	ATCGTGTTGTCCGGCAGGAAGTTCT
LT119	TTATTGCTTCTCTTGACCGTAGGAC	GTCCTACGGTCAAGAGAAGCAATAA	LT160	GTCGTGTATTCCGGACAGTACGTGG	CCACGTACTGTCCGGAATACACGAC
LT120	AACGGTTCGACCTTCTAATCCTATC	GATAGGATTAGAAGGTCGAACCGTT	LT161	CAAGCGTGACAAGTCCACGTATGACC	GGTCATACGTGGACTTGTCACGCTTG
LT121	ACTGACCTGTCGAGCTTAATATTCT	AGAATATTAAGCTCGACAGGTCAGT	LT162	CTTCGATCCGGTTGAGGTGGATATG	CATATCCACCTCAACCGGATCGAAG
LT122	TTGTCAGGTTACCAACTACTAAGGT	ACCTTAGTAGTTGGTAACCTGACAA	LT163	CACTGCTTCGCAGGCTGACGTTACT	AGTAACGTCAGCCTGCGAAGCAGTG
LT123	GGATGAATCGCTTGGTGTACCTCAT	ATGAGGTACACCAAGCGATTCATCC	LT164	AAGAAGTTCAGGAAGCGGTGATGCT	AGCATCACCGCTTCCTGAACTTCTT
LT124	TTGCTAGGCACTGATACATAACTCT	AGAGTTATGTATCAGTGCCTAGCAA	LT165	AAGAGGATGGTGCAGCAACCAACAA	TTGTTGGTTGCTGCACCATCCTCTT
			LT166	AAGTGAGAGGCACCTGTCAGATTGAG	CTCAATCTGACAGGTGCCTCTCACTT

使用 LT1 作为阳性对照，进行样点之间的标准化处理。使用 5′-FAM 标记反向互补序列。
在所有用于在微阵列上点样的标签的 5′ 端加上 14 个 Ts。

- **微阵列的构建**。标签寡核苷酸点样产生足量的微阵列，可供多次实验中使用。这些载玻片最多可保存三个月。
- **基因分型**。由 PCR 扩增和随后进行的 SBE 反应组成。基因分型耗时 1～2 天。
- **杂交**。基因分型的产物与微阵列杂交，并进行扫描。杂交过程耗时 1～2 天。
- **数据分析**。在微阵列数据经过质量控制和标准化之后，确定基因型。

SBE-TAGS 的设计与使用

- **标签寡核苷酸**：从包含少数重复单元的、序列已知的非相关基因组产生独特的序列标签，例如，噬菌体 λ 基因组。引物选取软件（例如，PRIMER 3.0；http://www.genome.wi.mit.edu）可以用来识别潜在的标签寡核苷酸。标签寡核苷酸的设计参数包括：

  长度（25～27 个碱基）；

  T_m（盐浓度 50 mmol/L 时，约 55℃）；

  自身相似性低（引物末端少于 3 个碱基，内部碱基少于 8 个）；

  G + C 含量（40%～60%），并且，连续相同的碱基不超过两个。

  在这个集合中，挑选出彼此相似性低的寡核苷酸。通过 BLAST（http://www.ncbi.nlm.nih.gov/blast/blast.cgi）来确保寡核苷酸的序列与作为研究对象的基因组之间不同源。表 6-3 中列举了一组适用于人类和小鼠基因组的标签。
- **阳性对照寡核苷酸**：设计并合成一个与其中一条标签反向互补的序列作为阳性对照寡核苷酸，并在其 5′端连接 FAM 标记。如使用表 6-3 中的标签，使用序列 5′-FAM-TACATATCA-CAACGTGCGTGGAGGC 作为阳性对照寡核苷酸。
- **向标签寡核苷酸上添加连接序列**：向标签序列的 5′端加上 15（dT）残基，一旦其附着到阵列上，可促进它与附着的寡核苷酸之间杂交。
- **PCR 引物设计**：使用程序 PRIMER3.0 来挑选引物，可从 www.genome.wi.mit.edu 获得。参数应该设置为：T_m：57℃～63℃；CG 含量：20%～80%；引物长度：15～30 个核苷酸；以及产物最大长度为 120 个核苷酸。扩增子应该尽可能的小，这一点非常重要，这使得全部 PCR 产物的扩增效率接近一致。
- **SBE 序列的选择**：选择紧靠 SNP 5′端的一段长度为 15～30 个寡核苷酸，解链温度为 50～60℃的序列。SBE 引物序列不能与任何其他的 SNP 发生重叠。在这里介绍的方案中，不要选择腺嘌呤和胞嘧啶作为将要掺入碱基的多态性位点，因为这两种碱基会导致带有相同染料（TAMRA）的 ddNTP 的掺入。对于这些 SNP，使用反义链，也就是说，把它们作为 G/T，而非 A/C 多态性进行检查。
- **Tag-SBE 引物的设计**：使用迭代过程来选择标签与 SBE 序列的组合。每条 SBE 序列与一个标签序列的反向互补序列随机配对，产生在 5′端包含标签的混合的 Tag-SBE 引物序列。该引物的总长度为 41～52 个碱基。然后用下面给出的算法来预测所设想的引物是否会以使其不适用于 SNP 基因分型的方式发生自我折叠。该算法通过比较引物 3′端的序列与位于引物主体部分的潜在互补序列，给每个潜在的内部引导（internal priming）位点评分。引物的末端碱基必须与互补序列的末端碱基互补，并且在序列中两个互补片段的发夹环（hairpin turn）处至少有一个碱基，才可将其定义为潜在的互补序列。对序列的两段片段进行比较，通过以下算法进行自我折叠评分：

  自我折叠评分＝（匹配碱基数目）－1.7×(错配碱基数目)

  如果倒数第二个碱基错配，再减去 0.8

  自我折叠指数大于 3 的引物需要通过将 SBE 引物与一条不同的标签配对来重新设计。如果是 SBE 引物自身发生自我折叠，如果可能，应该在相反的 DNA 链上设计 SBE 引物。

- **构建兼容的 SBE 引物合并集合**：Tag-SBE 引物按照约 30 个位点一组合并在一起。为了防止在多重 SBE 反应中出现交叉引导（cross-priming），对 SBE 引物进行分组，使得在每个合并的集合中，应用上面介绍的算法时，没有一对引物给出的自我折叠评分大于 8。交叉引导的截止评分要比自我折叠高，因为产生分子内折叠需要的互补程度要比形成稳定的分子间杂交体所需要的互补程度低。在一组寡核苷酸中，成员之间的交叉引导可以通过不添加 DNA 模板的多重 SBE 反应来检测［也就是说，在没有任何 PCR 模板存在时，使用 SBE 反应产物（见步骤 14）］。表现出明显信号的引物需要使用 SBE 引物在相反的 DNA 链上进行重新设计。或者，也可以改变合并物中的成分，但是重新设计不合适的引物通常要简便一些。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

缓冲液，溶液和试剂

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

$MgCl_2$（25 mmol/L）

SBE 缓冲液

　　500 mmol/L Tris（pH 9.5）

　　20 mmol/L $MgCl_2$

SDS（20%）〈!〉

20× SSC

玻片封闭缓冲液和试剂

　　琥珀酸酐（Aldrich）〈!〉

　　1-甲基-2-吡咯烷酮（Aldrich）〈!〉

　　硼酸钠（1 mol/L，pH 8.0）

　　乙醇（95%）

玻片清洗缓冲液

　　2× SSC/0.1% SDS

　　2× SSC

　　0.2× SSC

酶和缓冲液

AmpliTaq Gold（5 单位/μl）（Applied Biosystems）

外切核酸酶Ⅰ（Amersham Biosciences）

10× GeneAmp PCR 缓冲液Ⅱ（4306894，Applied Biosystems）

SAP/Exo 缓冲液

　　500 mmol/L Tris-HCl（pH 8.5）

50 mmol/L $MgCl_2$

甲壳碱性磷酸酶 (Shrimp alkaline phosphatese) (Roche)

Thermo Sequenase 聚合酶 (32 单位/μl) (Amersham Biosciences)

核酸与寡核苷酸

噬菌体 λ DNA (500 μg/ml) (New England Biolabs)

dNTP 溶液，包含全部四种 dNTP，每种浓度均为 10 mmol/L (pH 8.0，Pharmacia)

荧光标记的 ddNTP (定制的，NEN Life Science Products)

羧基四甲基罗丹明 (TAMRA) -ddATP

羧基四甲基罗丹明 (TAMRA) -ddCTP

花青素苷 5 (Cy5) -ddGTP

羧基-*X*-罗丹明 (ROX) -ddUTP

两种碱基使用相同荧光标记是由于本方案中所使用的扫描仪 (ScanArray 5000) 只能检测四个不同的波长 (四通道和四个滤色块)。一个通道用于 FAM 对照，剩下三个通道/染料用于基因分型。根据经验确定实验中最有效的染料—碱基组合。

基因组 DNA

与常规 PCR 相同的制备方法，稀释到 5 ng/μl。

抽提的基因组 DNA 样品应该具有高的质量，建议使用来自 Gentra Systems，Inc. 的商品化 DNA 抽提试剂盒。

410

寡核苷酸 (GIBCO Custom Primers，常规合成)：

- 用于微阵列点样的标签 (关于所使用标签序列的示例，见表 6-3) 用蒸馏水溶解，浓度为 200 μmol/L。阳性对照寡核苷酸用蒸馏水溶解，浓度为 100 μmol/L 和 20 μmol/L。
- 5-羧基荧光素 (FAM) 标记的阳性对照寡核苷酸 (例如，来自表 6-3，反向互补序列：5′-FAM-TACATATCACAACGTGCGTGGAGGC)，用蒸馏水稀释到5 μmol/L。
- 未修饰的 PCR 引物，用于单重 PCR，(请见本方案中 **SBE-TAGS 的设计与使用**信息栏)。每个 SNP 的正向和反向引物在溶液中成对存在，且各自在水中的最终浓度为 1 μmol/L。
- Tag-SBE 引物 (见本方案中 **SBE-TAGS 的设计与使用**信息栏)。引物在溶液中合并，足够分析约 30 个 SNP，每条引物的最终浓度为 1 μmol/L。

鲑鱼睾丸 DNA (10 μg/ml) (Sigma)

专用设备

天平 (用于称量和混合玻片封闭溶液)

盖玻片 (PMM-3，Lab Scientific，Livingston，New Jersey)

干燥箱 (VWR Scientific Products)

镊子 (尖头与钝头的)，用于夹持盖玻片

玻璃瓶（500 ml）
玻璃记号笔（金刚石笔），用于对玻片进行编号（VWR Scientific Products）
量筒（250 ml 和 50 ml）
加热槽，设定为 95 ℃
杂交盒（根据 P. O. Brown 网站上的设计图修改，http：//cmgm. stanford. edu/pbrown/mguide/HybChamber. pdf，或者从 Eastern Tool，Cambridge，Massachusetts 定制。）
湿度计，用于监测点样仪的点样室内湿度（VWR Scientific Products）
微阵列分析软件（如 ArrayVision，Imaging Research Inc.，Ontario，Canada）
微阵列扫描系统（如配有外置式氩离子激光的 ScanArray 5000，Packard Biochip Technologies，Billerica，Massachusetts）
微阵列点样仪（GMS417 Arrayer，Affymetrix）

将增湿器连接到点样仪，并在点样仪中放置一个湿度计，用来监测点样室内的湿度。配有湿度调节器以及内置湿度计来监测舱内湿度。

微量滴定板和多孔板密封膜

按照 Affymetrix 的推荐使用深孔的 96 或 384 孔板进行点样（任何来自 Marsh 或 MJ Research，或等同的供应商的 V 形底部 96 或 384 孔板均可适用）；使用多孔板密封膜（Marsh 可提供）或者 Microseal A 密封膜（MJ Research 提供）.

轨道振荡器（VWR Scientific Products）
烘箱，设定为 50℃
平板清洗器，用于多重 PCR 中 Dynabeads 磁珠的清洗步骤（SLT Lab Instruments，根据定制进行修改，加装 96 孔磁圈底座）
PRIMER3.0 引物设计软件（http：//www. genome. wi. nlit. edu）
玻片干燥台（VWR Scientific Products）
玻片架（Thermo Hybaid，Franklin，Massachusetts）

将金属玻片架手柄的一处弯曲拉直，以适合放入离心机的载样托盘。

玻片染色皿组件（VWR Scientific Products）
玻片保存盒（避光的）（VWR Scientific Products）
玻片，包被有聚-L-赖氨酸（Lab Scientific，Livingston，New Jercy）或者硅烷（Telechem，Sunnyvale，California）

这两种玻片具有可比性，但是就作者的经验而言，来自 Telechem 的硅烷化修饰玻片更为一致。

Stratalinker 紫外交联仪（Stratagene）
台式离心机（VWR Scientific Products）
热循环仪

热循环仪根据扩增方案进行编程（MJ Research DNA Engine 或 Terad，或者 Perkin Elmer 2400、9600 或 9700），最好配有热盖。

水浴，预设为 55℃

方法

制备用于 SBE 的玻片微阵列

将标签寡核苷酸点样于修饰过的玻片表面来制备用于 SBE 的微阵列。第 2 章，方案 2 中讨论的许多问题，如各种自动点样设备与微阵列玻片扫描仪，都与本方案相关。

1. 制备加样板：

   a. 按以下配比为加样板各孔准备混合物：

加样板成分	体积/加样孔/μl
阳性对照标签寡核苷酸（100 μmol/L）	3.0
20×SSC	4.5
dH_2O	7.5
总体积	15.0

   b. 向 96 孔微量滴定板（加样板）的每孔中加入 15 μl 该混合物。

   c. 向每孔中加入 15 μl 特异性标签（200 μmol/L），忽略加样板上 6 个分散的孔。

   d. 向这 6 个孔中再加入 15 μl 阳性对照标签（20 μmol/L）。这些孔单纯作为阳性对照。

   此时加样板上每孔中应该含有 30 μl 的液体。按照以下说明，用这么多的材料，Affymetrix 417 Arrayer 可以点 168 张玻片。

2. 在增湿的点样室内（湿度 40%），用 GMS417 微阵列点样仪将每种寡核苷酸点印在玻片上四个相邻的位置上，如果阵列中样点在空间上分散很好时，也可以重复成对点样，在聚-L-赖氨酸包被的载玻片（或硅烷处理的载玻片）上进行点样，速度为每个样点点印 3 次。

3. 因为此时点样的微阵列在玻片表面清晰可见，准备模板以使得微阵列在杂交过程中容易被识别（步骤 15～16）。在一张纸上，画出玻片的轮廓和微阵列在其中的位置。确保在每个模板上标明玻片磨砂边沿的位置。

处理和封闭点样后的玻片

4. 将玻片置于玻片架上，点样的一面朝下，在 55℃的水浴上方放置约 30 s，使之润湿。然后将点样的一面朝上，放置在玻片干燥台上（预热至 100 ℃）约 2 s，使玻片快速干燥。

5. 将玻片放入紫外交联仪中，点样的一面朝上，在 60 mJ 下使点样的寡核苷酸与玻片表面发生交联。

6. 按照以下步骤对玻片进行封闭：

   a. 将 3.75 g 琥珀酸酐溶解于 236.25 ml 1-甲基-2-吡咯烷酮中配制封闭溶液。将玻片放入玻片染色组件中的移动式玻片托盘内，并将容器放入干净的染色皿中。

b. 迅速将 26.25 ml 1mol/L 的硼酸钠溶液加到琥珀酸酐的 1-甲基-2-吡咯烷酮溶液中。立即涡旋混合溶液，并将封闭溶液即刻倒于玻片上。盖上染色皿，置于轨道振荡器上，以约 75 r/min 的速度振荡 15 min。

c. 在一个新的染色皿内装入 95%乙醇。在轨道振荡器（步骤 b）上处理 15min 之后，从封闭溶液中取出玻璃托盘并转移到装有乙醇的染色皿中。轻轻地上下移动浸泡托盘 1 min 来清洗玻片。

d. 将玻片转移到经过改制的金属玻片架中（玻片架经过修改，将手柄的一处弯曲拉直），并在室温下以 500 r/min（低速）离心约 5～10 min 以干燥玻片（玻片架保持平衡）。

7. 用玻璃记号笔对玻片进行编号，并置于避光盒内。

在干燥空气中，玻片可以保存至少 3 个月。

微阵列的质量控制

8. 从每个批次的玻片中取出一片，与 FAM 标记的反向互补序列阳性对照进行杂交（步骤 16～19）。按照步骤 16 中的说明操作，除了用 13.3 μl 的水代替 SBE 产物（步骤 16a）。

测试各批次玻片中的哪一张玻片没有问题——尽管一些研究人员宁愿弃去开始的 1 或 2 张玻片。

9. 扫描并分析该玻片（步骤 21～24）。只有当阳性对照得到满意的信号并且各点之间的信号强度差异较小时，方可使用这批玻片。

进行 PCR 和 SBE 反应

PCR 反应可以先单个逐一进行，之后再将产物混合在一起，或者，也可以进行多重反应。这里给出了单重的方案，在本方案最后**替代方案：为 SBE 进行多重扩增反应**的信息栏中给出了多重扩增的方法。单重 PCR 反应具有许多优点：它不需要相容的 PCR 引物集合，而且扩增更强，清除方法也比较简单。但是，多重 PCR 节省模板用量，并且通过较少次数的扩增反应就可以完成任务。

10. 将实验基因组 DNA 用水稀释至 5 ng/μl。为在步骤 11b 中要用的每份样品准备一支标记好的离心管。

11. 按照下述步骤在微量滴定板中建立单重扩增反应：

a. 将与给定的 SNP 对应的 PCR 引物对加入到 PCR 多孔板的各个孔中（总体积为 5 μl，其中每种引物的浓度为 1 μmol/L）。

b. 为每个实验基因组 DNA 配制 PCR 预混试剂。计算出每孔中要扩增的混合物的量，再增加总体积的 10%的冗余量，以补偿移液时的损失。

c. 向多孔板中每个装有特定基因组 DNA 的孔中加入 10 μl 预混试剂。

有时在镁离子的其他浓度下做一些分析效果会更好，但是，这个反应中通常不对镁离子浓度进行优化。可能一个更为重要的变量是 dNTP 的浓度，因为过量的 dNTP 会抑制接下来的 SAP/Exo 的清除步骤。

由于预混试剂中含有 DNA，所以其中不能包含任何具有生物学意义的个体阴性

对照，唯一可能的阴性对照就是整个多孔板中的全部引物，要用水来替代预混试剂中的 DNA。

试　剂	反应体积/μl
10×PCR 缓冲液Ⅱ	1.5
25 mmol/L $MgCl_2$	0.9
10 mmol/L dNTP 溶液	0.15
5 单位/μl Ampli Taq Gold	0.15
5 ng/μl DNA	2.5
dH_2O	4.8
总体积	10.00

12. 将微孔板放入热循环仪中，按照下列的变性、复性、聚合时间和温度进行扩增反应：

循环次数	变　性	复　性	聚合（延伸）
1	96℃，10 min		
35	96℃，30 s	50 ℃，1 min	72℃，1 min

这些时间是根据本方案中所介绍的专用设备和反应而进行优化的，需要对这些条件进行修改以适合其他类型的设备和反应。得到的 PCR 产物的平均大小为 50～150 个碱基。

13. 从热循环仪中取出微孔板，每个样品的反应取 1 μl 合并在一起，得到一个 PCR 产物集合。合并时，只能是扩增自相同个体的位点，并且是已知 SBE 引物之间相容的那些产物（见本方案导言中**SBE-TAGS 的设计与使用**信息栏里关于构建相容的 SBE 引物集合的讨论）。每个集合中应该包含约 30 个样品的 PCR 产物。留下 10 μl 合并的产物用于下面的清除步骤（步骤 14）。剩余的 PCR 产物可在 4℃下保存数月。
14. 用甲壳碱性磷酸酶消化去除过量的 dNTP，并用核酸外切酶Ⅰ消化去除多余的引物，以纯化 PCR 产物。

414

a. 为每个合并的 PCR 产物集合制备以下反应混合物：

试　剂	反应体积/μl
1 单位/μl 甲壳碱性磷酸酶	2.0
20 单位/μl 核酸外切酶Ⅰ	0.2
SAP/Exo 缓冲液	1.5
PCR 产物（来自步骤 13）	10.0
dH_2O	6.3
总体积	20.0

b. 将反应混合物在 37 ℃下温浴 45 min，然后再在 96℃下温浴 15 min，使酶失活。得到 10 μl 清除后的产物用于下面的 SBE 反应（步骤 15）。剩余材料可以在 4 ℃下保存数月。

15. 进行 SBE 反应：

a. 为每一个 PCR 产物集合制备 SBE 反应混合物：

试　剂	反应体积/μl
SBE 缓冲液	1.0
Tag-SBE 引物集合（每条引物 1 μmol/L）*	1.0
32 单位/μl Thermo Sequenase	0.06
100 μmol/L TAMRA-ddATP	0.05
100 μmol/L TAMRA-ddCTP	0.05
100 μmol/L Cy5-ddGTP	0.05
100 μmol/L ROX-ddUTP	0.05
dH_2O	2.74
纯化的 PCR 产物（来自本方案步骤 14 或替代方案中步骤 4）	10.00
总体积	15.00

每个 SBE 反应包含一个阴性对照（不含 PCR 产物）

b. 将 SBE 反应放入热循环仪中，根据下列反应条件进行 SBE：

循环次数	变　性	复　性	延　伸
30	96℃，30s	50 ℃，15 s	60℃，1 min

Tag-SBE 引物的摩尔数要超过 PCR 模板。在每次循环反应中，部分引物被延伸，并在接下来的反应中与未修饰的引物产生竞争。30 个循环的扩增可以使微阵列杂交得到较好的信号强度。

c. 从热循环仪中取出微孔板，如果要在第二天继续本方案，将 SBE 产物在 4℃下保存过夜，否则继续进行步骤 16（杂交）。

SBE-TAGS 产物的微阵列杂交

将来自步骤 15 的标记的 SBE 产物加到微阵列上进行杂交反应。许多步骤与第 3 章的方案 20 类似，其中还配有逐步的图解以帮助初次实验者。

415

16. 在离心管中合并以下试剂制备杂交缓冲液：

试　剂	12 张玻片需要的总体积/μl
20×SSC	18.0
20% SDS	0.9
鲑鱼睾丸 DNA（10 μg/ml）	0.9
FAM 标记的阳性对照寡核苷酸（5 μmol/L）	3.0
dH_2O	67.2
总体积	90.0

为防止 SDS 沉淀，将装有溶液的离心管置于 95℃的加热槽中。在进行下一步反应之前再将试管从加热槽中取出。不要将试管置于冰上。

a. 向 6.7 μl 杂交缓冲液中加入 13.3 μl 单个的 SBE 产物（来自步骤 15）。将全部体积的混合溶液（共 20 μl）直接加到微阵列上（为了准确定位微阵列，将玻片置于步骤 3 所准备的模板上）。用镊子除去盖玻片上的封膜并小心地盖在杂交混合物上，仔细操作以防止下方出现气泡。用镊子轻敲盖玻片使气泡从其边缘排出。对每个要进行基因分型的 DNA 重复本步骤。

b. 将载玻片置于密封的小杂交室内，并在玻片的磨砂边沿加上 55 μl 水，以限制液体蒸发。

17. 将微阵列在 50℃下杂交 4 h。

18. 清洗微阵列：

a. 用玻片染色皿，在室温下用 2×SSC/0.1% SDS 清洗微阵列 5 min（经常振荡），然后用 2×SSC 快速漂洗 3 次，最后用 0.2×SSC 漂洗 1 min。

b. 室温下，将玻片以 500 r/min 的速度离心 15 min 除去残留的洗涤液。

19. 将干燥的玻片保存在避光的玻片盒中，并在几天之内进行扫描。

微阵列的扫描

本方案中所使用的多种荧光基团存在重叠的激发和发射光谱，导致了少量的“交叉干扰”，即来自一种染料的信号在一个以上的通道中被检测到。该问题依赖于所使用的仪器和具体设置，可以用矩阵进行校正，该矩阵是根据经验为每个扫描方案确定的。

20. 确定 3×3 的校正矩阵（表 6-4，表 6-5）

a. 分别使用三种滤色块，扫描带有三种“纯染料”的样品，其中每个样品含有三种染料中的一种（TAMRA、ROX 和 Cy5）。

b. 对每个纯染料样品，将每个滤色块的数据表示为样品中的染料所用的滤色块得到的信号强度的比例。

c. 将这些数据排列成 3×3 的矩阵，且对角元素均为 1。校正矩阵很简单，就是该 3×3 矩阵的逆矩阵。

表 6-4 反映各种染料之间交叉干扰程度的矩阵

加入的染料	通 道		
	TAMRA	ROX	Cy5
TAMRA	1	0.26	0.29
ROX	0.06	1	0.02
Cy5	0.03	0.06	1

表 6-5 表 6-4 中数值的校正或逆矩阵

加入的染料	通 道		
	TAMRA	ROX	Cy5
TAMRA	1.023453086	−2.53515888	−0.290777954
ROX	−0.60654832	1.016085402	−0.001023174
Cy5	−0.02573221	−0.051794878	1.008371708

通过激发带有 TAMRA、ROX 或 Cy5 标记的 SBE 产物（行）的玻片来评价干扰的程度，用适合每种染料的通道来读取数据（列）。数值表明了对于一种给定的染料，不正确通道中的信号强度占正确通道中信号强度的比例。

21. 用类似于 ScanArray 5000 这样带有外置氩离子激光器的微阵列扫描系统，以 20 μm 的分辨率对阵列进行扫描。针对这种特定的仪器，用以下设置：
    FITC：激光 100%，PMT 90%
    TAMRA：激光 80%，PMT 80%
    ROX：激光 90%，PMT 90%
    Cy5：激光 80%，PMT 80%

FITC 激光用来检测阳性对照寡核苷酸中 FAM 的荧光信号。

数据分析

关于数据分析的详细讨论，见第 7 章。

22. 确定在总信号中由其他两种相关染料产生的信号分别占的相对比例。
    a. 使用微阵列分析软件包，如 ArrayVision（Imaging Research Inc.）减去每个通道的局部背景。在每个样点上，根据 FAM 标记阳性对照寡核苷酸的强度将信号进行标准化。如果阳性对照信号非常微弱（小于玻片平均值的 10%），就将该点当成缺失处理。
    b. 将来自 TAMRA、ROX 和 Cy5 的经过标准化的数据乘以校正矩阵（来自步骤 20），得到校正的“纯染料”数值。为了校正染料终止物的不同亮度和掺入效率，对于掺入了 C 的 SNP，将矩阵修正后的 Cy5 信号乘以 0.5，矩阵修正后的 TAMRA 信号乘以 1.5。这些修正因子可能需要根据样品、仪器和所使用的参数进行相应的调整。

根据对多种 SNP 的已知基因分型的样品进行检测的结果来决定是否需要进一步的调整，实

验中杂合子的比值在平均后的期望值应为 50：50。

23. 确定数据的质量是否足够用于进行基因分型。剔除整体杂交信号低的玻片。
    a. 玻片上的数据点的平均信号应该超过阈值，该阈值定义为所有样点的平均信号强度的 1/10，包含阳性对照和 SNP 检查的信号强度。
    b. 确定四个重复样点的平均信号大于全部玻片中这些样点平均信号的 25%。
    c. 忽略信号弱的个别样点（如可能由于杂交过程中微小的气泡所造成的），用剩余的重复样点来确定基因型。
    d. 所有重复样点中的基因型应该一致。

24. 按照以下方法给每个样点计算基因分型分值：

    基因分型分值＝lg［（标记等位基因 A 的染料所占比例＋0.01）/（标记等位基因 B 的染料所占比例＋0.01）］

按照 SNP 来对基因分型分值进行聚类。对于每个 SNP，基因分型分值应该属于三种基因型的聚类之中：

纯合子 A（AA）个体的分值通常小于－1。

纯合子 B（BB）个体的分值通常大于＋1。

杂合子（AB）个体的分值总是介于－1 与＋1 之间。

当分值为±1 时，表示两种染料的比值为 10：1。

对于聚类分割不清晰的 SNP，不能进行基因分型。要注意聚类方法在数据点很多时效果很好，因此，在对每个 SNP 进行基因分型时，建议同时分析尽可能多的样本。

疑难问题解答

对于任何一种基因分型方法，强烈建议引入部分重复的基因型或者相关的样本（用来检查遗传性误差），以保证低的误差率。对于数据点聚类不清晰的情况，建议使用目测或半自动的方法来筛查 SNP。

418

替代方案：为 SBE 进行多重扩增反应

当模板 DNA 的来源受到限制的时候，通常进行基因组 DNA 的多重而不是单重扩增。使用单重 PCR 对一个位点进行扩增和基因分型所需要的 DNA 数量足够使用多重 PCR 对 50 个位点进行扩增和基因分型。但是，多重 PCR 反应和随后的清除方案与单重 PCR 相比却要花费更多的人力。

在多重扩增方案中，使用 5′端加尾的引物进行第一次扩增。使用生物素标记的对应于加尾部分的引物进行第二次扩增。由于许多位点在反应中的竞争，该策略用来增加 PCR 产物的产量。

多重反应中高浓度的引物和 dNTP 的存在，需要使用更加高效的清除步骤。由于第二次的 PCR 引物是生物素化的，可以使用链亲和素包被的磁珠来结合第二次扩增产生的生物素化产物。

多重方案遵循主要方案的顺序，例外之处在相应步骤的下方注释。

附加材料

生物素化的引物，为多重 PCR 的第二轮而修改

T3 引物：5′-生物素-AATTAACCCTCACTAAAGGGAGA

T5 引物：5′-生物素-TAATACGACTCACTATAGGGAGA

这些引物，在 5′端带有生物素标记，与附加到第一轮扩增引物上的噬菌体 T3 和 T7 序列相同。生物素标记的引物在溶液中成对存在，每种引物浓度为 20 μmol/L，溶解在 H_2O 或 0.1×TE 中。

Dynabeads 磁珠，M-280 链亲和素（10 μg/ml）（Dynal Biotech Inc.，Lake Success，New York）

Dynabeads 结合缓冲液

TE（pH 8.0）

2 mol/L NaCl

Dynabeads 清洗缓冲液

70%乙醇

3 mmol/L Tris（pH 8.0）

0.3 mmol/L EDTA（pH 8.0）

微阵列，用标签寡核苷酸点样，制备方法按照主要方案中的步骤 1～9 进行，用于杂交。

PCR 引物，为第一轮多重 PCR 而修改

附加到正向引物 5′端的 T3“加尾”序列：

5′-AATTAACCCTCACTAAAGGGAGA

附加到反向引物 5′端的 T7“加尾”序列：

5′-TAATACGACTCACTATAGGGAGA

引物合并到包含约 30 个 SNP 检查的溶液中，每种引物的终浓度为 1 μmol/L。

方法

用于多重扩增的标签和 Tag-SBE 引物的设计

1. 根据主要方案的导言最后 **SBE-TAGS 的设计与使用**信息栏中给出的建议，设计带有附加接头序列的标签寡核苷酸。
2. 设计和修改用于多重扩增反应的 PCR 引物：
   a. 使用程序 PRIMER3.0 来选择引物序列，该程序可从 www.genome.wi.mit.edu 得到。参数应该设置为：T_m：57～63℃；C+G 含量：20%～80%；引物长度：15～30 个核苷酸；以及最大产物长度为 120 个核苷酸。扩增子应尽可能地短，以便于同等地扩增。
   b. 修改和汇集多重 PCR 引物：
      i. 将 T3（AATTAACCCTCACTAAAGGGAGA）尾添加到每个正向引物的 5′端，T7（TAATACGACTCACTATAGGGAGA）添加到每个反向引物的 5′端，以此来修改第一次扩增中使用的引物。

ii. 在第二次扩增中，配制浓度为 20 μmol/L 的 5′-生物素化-T3（5′-biotin-AATTAACCCTCACTAAAGGGAGA）和 5′-生物素化-T7（5′-biotin-TAATACGACTCACTCATAGGGAGA）溶液。

c. 决定相容性位点的 PCR 集合时，要确保集合中的位点之间不存在大的同源区域。使用主要方案导言中 SBE-TAGS 的设计与使用信息栏中给出的自我折叠算法，但是要将对应的 SBE 引物与全体 PCR 产物进行比较，而不仅仅是和集合中的 SBE 引物进行比较。

3. 根据主要方案的导言最后 **SBE-TAGS 的设计与使用**信息栏中给出的建议，选择 SBE 序列和设计 Tag-SBE 引物来构建 SBE 引物的相容性集合。

多重 PCR 和 SBE 反应

4. 进行多重 PCR：

a. 将基因组 DNA 稀释为 2 ng/μl。

b. 按照以下方法在微孔板中建立第一次扩增反应：

i. 向每孔中加入 1.25 μl 合并的 PCR 引物，依次加入所有对应于特定基因组 DNA 的 PCR 引物合并物。

ii. 向每孔中加入 1.25 μl（2 ng/μl）基因组 DNA。

iii. 为每个 PCR 集合制备扩增预混试剂，根据 PCR 集合的数量计算需要的总体积，增加 10%的冗余。

试　剂	反应体积/μl
10×PCR 缓冲液 Ⅱ	1.25
25 mmol/L $MgCl_2$	1.042
10 mmol/L dNTP 溶液	0.625
5 单位/μl AmpliTaq Gold	0.5
dH_2O	6.583
总体积	10.0

iv. 向微孔板的每孔中加入 10 μl PCR 混合物。

c. 将微孔板放入热循环仪中，并按照下列的变性、复性、聚合时间与温度进行扩增反应：

循环次数	变　性	复　性	延　伸
1	95℃，9 min		
31	95℃，30 s	55℃，30 s	72℃，30 s
1			72℃，5 min

这些设定时间只是为本方案所介绍的特定设备和反应而优化的，因此需要对条件进行修改以适合其他类型的设备和反应。

d. 建立第二轮扩增反应

i. 向一个新的微孔板的每孔中加入 3 μl 第一轮 PCR 的产物。

ii. 配制第二次扩增的混合物，根据第一次扩增产物的数量计算需要的总体积，加上 10%的冗余。

试　剂	反应体积/μl
10×PCR 缓冲液 Ⅱ	3.0
25 mmol/L $MgCl_2$	4.8
10 mmol/L dNTP 溶液	1.2
5 单位/μl Ampli Taq Gold	0.6
生物素标记的 T3 与 T7 引物，每种 20 mmol/L	1.2
dH_2O	16.2
总体积	27.0

iii. 在新微孔板的每孔中加入 27 μl 扩增混合物。

e. 将微孔板放入热循环仪中，并按照下列的变性、复性、聚合时间与温度进行扩增反应：

循环次数	变　性	复　性	延　伸
1	95℃，8min		
32	95℃，30s	55℃，1min，30s	72℃，30s
1			72℃，7min

这些时间只是为本方案中所介绍的特定设备和反应而优化的，因此需要对条件进行修改以适合其他类型的设备和反应。

5. 使用链亲和素包被的 Dynabeads 磁珠纯化 PCR 产物，从扩增产物中除去多余的 dNTP 和引物：

a. 对于每种第二次 PCR 产物，在圆底微孔板中混合以下试剂：

试　剂	反应体积/μl
Dynabeads 磁珠	5.0
Dynabeads 结合缓冲液	12.0
第二轮扩增产物	7.0
总体积	24.0

b. 室温下，将结合反应物在磁性板上温浴 30min。

c. 用 SLT 改进的平板清洗器，在 Dynabeads 清洗缓冲液中清洗结合反应物 3 次，然后除去清洗缓冲液。

清洗也可以手工完成。用移液器缓慢地上下吸打清洗混合液，小心地将吸头放在板孔中央而不要接触到孔壁上的一圈 Dynabeads 磁珠。

d. 将磁珠置于空气中干燥数分钟，然后将每组带有 PCR 产物的磁珠重悬于 10 μl 水中。

6. 回到主要方案中的步骤 14，并将得到的产物用于 SBE 反应，确保 SBE 集合中的位点与 PCR 集合中的位点相对应。按照主要方案（步骤 16～24）继续进行杂交、扫描和分析。

421

方案 7

Affymetrix 标签阵列用于 SBE SNP 基因分型

Thomas Ryder，Xiaohua Huang，Nila Patil，Jian-Bing Fan（昂飞公司，圣塔克拉拉，加利福尼亚 95051）

本方案介绍了应用 Affymetrix GenFlex 阵列对单碱基多态（single-base polymorphisms，SNP）进行 Affymetrix 阵列基因分型。在方案 9 中介绍的人类 SNP（HuSNP）阵列使用了特定的、与特异性 SNP 两侧序列互补的探针。因此，用 HuSNP 阵列所能检测的 SNP 范围是由阵列的组成所决定的。与此相比，GenFlex 阵列使用通用的 20 聚体寡核苷酸标签探针，在原理上，借助于具有双重功能的标签引物进行单碱基延伸（SBE）反应，可以检测任意的 SNP。每条标签引物的 5′端与阵列中的标签探针反向互补，而 3′端对应于给定的 SNP 的紧靠 5′端的序列（图 6-9）。标签引物的 3′碱基终止于 SNP 位点的前一个碱基处。

在 SBE 反应中（图 6-9），每个 SNP 片段首先被 SNP 位点侧翼的引物扩增。这些扩增反应通常是多重的，因此每次反应扩增 20～30 个位点。然后，在 *Taq* 聚合酶或 Thermo Sequenase 聚合酶的催化下进行 SBE 反应，这个反应也是多重的。标签引物与 PCR 产物复性，对应于双等位基因 SNP 多态性的两种可能碱基之一，用双脱氧核苷酸三磷酸盐掺入。由于另一种可选择的碱基上标记有不同的荧光基团或半抗原，因此可以对掺入碱基的性质进行识别。使用链终止双脱氧核苷酸确保了只有单个碱基可以被掺入。然后将 SBE 合并，并与标签阵列进行杂交。使用激光扫描仪进行数据采集，并根据观察到的相对信号强度来确定基因分型结果。类似的方法已经被应用于使用预合成寡核苷酸点样的微阵列中（见方案 6）（Hirschhorn et al. 2000；Lindblad-Toh et al. 2000b）。

本方案中所使用的基本方法最初由 Fan 等（2000）提出，并成为商品化的 GenFlex 标签阵列的原型。Fan 等介绍了一种包含有 32 000 个特异性的完全匹配（perfect-match，PM）标签探针，和另外 32 000 个错配（mismatching，MM）寡核苷酸的阵列，其中错配探针与相关联的完全匹配探针在中间位置碱基上存在差别。在本书写作期间，GenFlex 标签阵列包含有 2050 PM 和 2050 MM 标签，在与标记的寡核苷酸杂交时得到高的 PM/MM

信号比。在这些探针当中，有 50 个具有代表性的探针被选作对照，留下 2000 个可用的探针用来检测 SNP 的性质。在每个杂交实验中，与对照标签探针互补的 50 个荧光标记寡核苷酸的集合，被加入到杂交混合液中，作为信号强度的对照。

制备用于 SBE 反应的双功能标签寡核苷酸时需要仔细考虑。在方案 6 导言里 **SBE-TAGS 的设计与使用**信息栏中介绍了它们设计上的一些注意事项，以及如何决定哪些标签可以在多重反应中被合并。 422

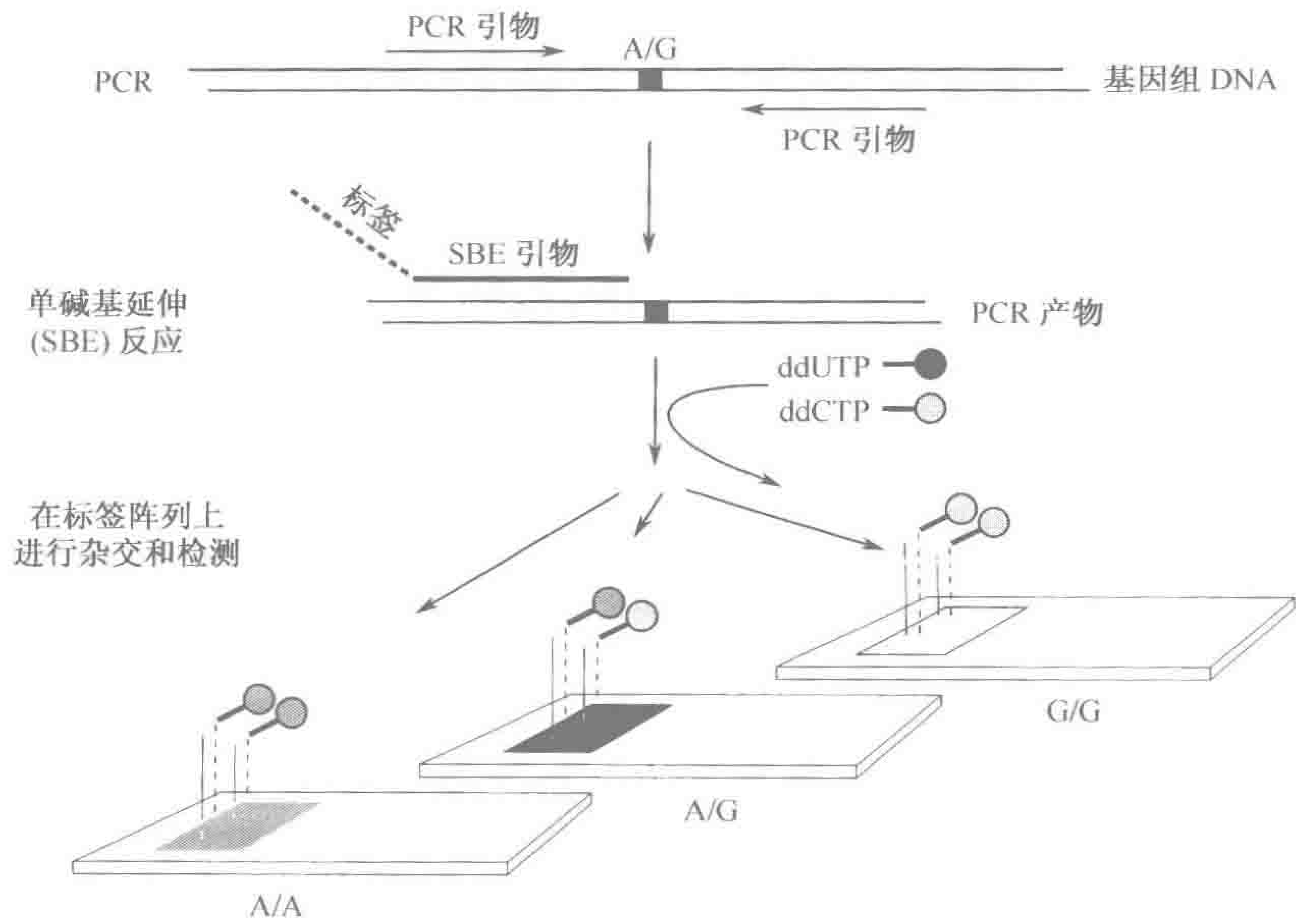


图 6-9　Tag-SBE 基因分型检查

设计带有特异性标记的引物用于从基因组 DNA 中扩增 SNP（Wang et al. 1998）。将带有相同的成对可变碱基（例如，A/G SNP）的 SNP 合并在一起。双链 PCR 产物用做 SBE 反应的模板。每个 SBE 引物嵌合有与阵列中特异标签互补的 5′端和与基因序列互补并且终止于多态性 SNP 位点之前一个碱基的 3′端。这样，每个 SNP 引物与阵列中的一个特异性标签唯一关联。在单个反应试管中，加入与多重标记对应的 SBE 引物，并在不同荧光基团标记的成对 ddNTP 存在下进行延伸；例如，在生物素标记的 ddUTP 和荧光素标记的 ddCTP 存在下，A/G 双等位基因标记进行延伸。标记的多重 SBE 反应产物合并在一起并与标签阵列杂交。如图所示，有三种杂交模式的结果，对应于三种基因型 AA、AG 和 GG（修改自 Fan et al. 2000.）

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

缓冲液、溶液和试剂

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

乙酰牛血清白蛋白（BSA）（50 mg/ml）（15561020，Invitrogen）

BSA（20 mg/ml）

乙醇，预冷至－20℃

糖原（100 μg/ml）（Boehringer Mannhein）

氯化锂（LiCl）（8 mol/L）

MES 缓冲液（0.5 mol/L，pH 6.7）

$MgCl_2$（5 mmol/L）

423

1×SSPE-T

用 0.001%的 Triton X-100 配制 1×SSPE。使用 0.2μm 过滤器过滤溶液。

6×SSPE-T

用 0.005%的 Triton X-100 配制 6×SSPE。使用 0.2μm 过滤器过滤溶液。

染色溶液

2.2 μg/ml 链亲和素 R-藻红蛋白

0.5 mg/ml 乙酰化 BSA

将试剂溶解于 6×SSPE-T。

链亲和素 R-藻红蛋白（1 mg/ml）（Molecular Probes）

氯化四甲胺（TMAC）（T 3411，Sigma）〈!〉

Triton X-100（T 9284，Sigma）

酶和酶缓冲液

AmpliTaq Gold（Applied Biosystems）

10×GeneAmp PCR 缓冲液Ⅱ（4306894，Applied Biosystems）

核酸外切酶Ⅰ（10 单位/μl）（Amersham Biosciences）

甲壳碱性磷酸酶（1 单位/μl）（Amersham Biosciences）

Thermo Sequenase 聚合酶（32 单位/μl）（Amersham Biosciences）

5×Thermo Sequenae 缓冲液（随酶提供）

核酸和寡核苷酸

生物素标记的 dNTP（New England Nuclear）

生物素-N6-ddUTP

生物素-11-acyclo-dCTP [（d）dCTP]

生物素-N6-ddATP

dNTP 溶液中包含全部四种 dNTP，每种的浓度为 25 mmol/L（dNTP 储存液，27-2035-01，Pharmacia Biotech）

荧光素标记的对照寡核苷酸（50 种合并在一起）（900302，Affymetrix）

荧光素标记的 dNTP（New England Nuclear）

荧光素-N6-ddGTP

荧光素-N6-ddCTP

荧光素-N6-ddATP

荧光素-N6-ddUTP

GenFlex Tag Array（5片包装）（900302，Affymetrix）

基因组DNA样品，稀释至20 ng/μl

PCR引物混合物

正向引物与反向引物定义了一段扩增子，这个扩增子覆盖了每个SNP位点。每个引物的浓度为1 μmol/L。

Tag-SBE引物混合物（每个SNP位点的TAG-SNP）

关于设计时的注意事项，见方案6导言中**SBE-TAGS的设计与使用**信息栏。每种引物的浓度为20 nmol/L。

专用设备

流体工作站（FS400，Affymetrix）

微阵列分析软件GeneChip（Affymetrix）

带有电动旋转架的可转动杂交箱

Microfuge微型离心机（5415C，Eppendorf）

S-300柱（Pharmacia）

424

扫描系统（定制，Affymetrix）

该系统对芯片上每个单元应该具有60～70像素的分辨率，并且设计有两个滤色镜，采集来自530 nm与570 nm波长的发射。

热循环仪

热循环仪按照需要的扩增方案进行编程（MJ Research DNA Engine或Tetrad，或者Perkin Elmer 2400、9600或9700），最好配有加热顶盖。

方法

多重扩增

由于SNP通常是双等位基因的，在一个SBE反应中，两种不同的标记足以对大多数SNP等位基因进行分辨（Marth et al. 2001）。为了便于后续的SBE反应，根据在多态性位点处具有相同碱基成分的SNP将位点扩增反应进行合并，例如，带有A或者G的全部SNP。合并的集合通常包含多达30个位点的扩增反应。

1. 按照下表配制用于多重位点特异性扩增的混合物，每个反应的最终体积为25 μl。

储存液	体积/μl	终浓度
10×缓冲液 Ⅱ	2.5	1×
25 mmol/L $MgCl_2$	5.0	5 mmol/L
25 mmol/L dNTP	1.0	1 mmol/L
PCR引物混合物（每种1 μmol/L）	2.5	0.1 μmol/L
基因组DNA（20 ng/μl）	2.5	50 ng
5单位/μl AmpliTaq Gold	0.4	2单位
ddH_2O	11.1	
最终体积/μl	25.0	

2. 将样品转移到热循环仪中，按照以下的变性、复性、聚合时间与温度进行扩增反应。

循环次数	变　性	复　性	延　伸
1	96℃，10 min		
40	94℃，30 s	57℃，40 s	72℃，90 s
1			72℃，10 min

将反应从热循环仪中取出，4℃下保存，备用于下一步骤。

3. 向步骤 1 的每个 PCR 产物（25 μl 反应）中加入 1 μl 核酸外切酶 I 以及 1 μl 甲壳碱性磷酸酶，在 37℃下保温 1 h 进行消化反应。然后 100℃下处理 15 min，使酶失活。

   必须对 PCR 产物进行酶处理，将残余的引物与 dNTP 分别进行降解和脱磷酸，因为它们影响 SBE 步骤。

4. 按照生产商（Pharmacia）的说明，将每个样品上样到 S-300 柱上，以纯化产物。

   将与待纯化样品等体积（约 25 μl）的溶液转移到收集管中，其中含有溶解于 ddH_2O 中的 PCR 扩增子，并且扩增子溶液已经过纯化除去了残留的 PCR 引物片段与核苷酸。这些样品可以直接用于 SBE 反应。

425　SBE 反应

5. 进行 SBE 反应：

   a. 按照下表所列配制 SBE 反应混合物。建立 6 个反应，每个反应对应于每种标记的与未标记的核苷酸对组合。

   SBE 反应在多重性的设置上与 PCR 反应具有同等水平，包含了最多 25 个待延伸的 SNP 混合物。关于引物设计以及确定哪些引物应该合并在一起的讨论，见方案 6 导言中 **SBE-TAGS 的设计与使用**的信息栏。

储 存 液	样品体积/μl	终浓度
经过扩增的模板（步骤 2）	6.0	
Tag-SBE 引物混合物（各 20 nmol/L）	2.5	各 1.5 nmol/L
5×Thermo Sequenase 缓冲液	6.6	1×
Flu-ddNTP（1 nmol/μl）	0.8	24 μmol/L
Bio-ddNTP（X nmol/μl）*	0.5	7.6 或 3.8 μmol/L
另外两种未标记的 ddNTP（1 nmol/μl）	各 0.3	各 10 mmol/L
6.4 单位/μl Thermo Sequenase Gold	0.4	2.56 单位
ddH_2O	15.6	
最终体积/μl	33.0	

* 当 Bio-ddNTP 为 Bio-ddUTP 时，浓度为 0.5 nmol/μl，或者 Bio-11-acyclo-dCTP（0.5 mmol/L），0.25 nmol/μl Bio-ddATP（0.25 mmol/L）。根据在给定的 SBE 集合中可能的等位基因型的信息来选择标记和未标记的 ddNTP；例如，均为 As 或 Gs 时将需要 Bio-ddATP，Flu-ddGTP 和未标记的 ddCTP 与 ddTTP。

b. 将反应转移到热循环仪中，并使用下列的变性和复性时间与温度进行 SBE 反应：

循环次数	变　性	复性与延伸
1	96℃，3 min	
45	94℃，25 s	58℃，11 s

这些时间只是为本方案中所介绍的特定设备和反应而优化的，可能需要对条件进行修改以适应其他类型的设备和反应。

c. 从热循环仪中取出反应产物，置于 4℃下保存直到进行以下步骤。

纯化

6. 将 6 个 SBE 反应（各 25 μl）合并在一起，并与 30 μl 浓度为 100 μ/ml 的糖原（Boehringer Mannhein）混合。用 18.75 μl 8mol/L 的 LiCl 和 11.25 μl 预冷的无水乙醇沉淀 SBE 反应产物。

   尽管不同的核苷酸使用了相同的荧光标记，但是全部 6 个延伸混合物（A/G、A/T、A/C、G/T、G/C 和 T/C）还是可以合并在一起并与一片阵列进行杂交，因为标签序列的性质和每个 SNP 位点所用的反应条件均是已知。

7. 将样品混匀，室温下以最大速度（16 000 *g*）离心 15 min 回收沉淀。弃去上清，将团块在 40℃下干燥 40 min，然后重悬于 33 μl 的 ddH_2O 中。

   SBE 产物可以直接用于杂交，如果不立即使用，可以冷冻保存。

426

SBE 产物的微阵列杂交

将来自步骤 7 的标记的 SBE 产物加到微阵列上进行杂交。参考第 3 章最后的疑难问题解答指南，其中提供了一组循序渐进的图解以帮助初次使用者。

8. 42℃下，用 6×SSPE-T 和 0.5 mg/ml 的 BSA 溶液对 GenFlex Tag Array 进行预杂交 15 min。
9. 混合以下试剂来配制杂交缓冲液：

储 存 液	体　积/μl	终浓度
5 mol/L TMAC	72.0	3 mol/L
0.5 mol/L MES 缓冲液	12.0	50 mmol/L
1% Triton X-100	1.2	0.01%
10 mg/ml 鲱鱼精 DNA	1.2	100 μg/ml
5 nmol/L Flu-c123 对照寡核苷酸	1.2	50 pmol/L
20 mg/ml BSA	3.0	100 μg/ml
SBE 产物（步骤 7）	29.4	
最终体积	120.0	

10. 将 SBE 样品置于杂交缓冲液（步骤 9）中，100℃下变性 10 min，然后在冰上快速冷却 2～5 min。

11. 除去 Tag Array 上的预杂交液，然后用移液器将全部杂交溶液（120 μl）加到 Tag 微阵列上。
12. 在转速为 40～50 r/min 的电动旋转架上，将 Tag Array 与 SBE-标记的产物在 120 μl 的杂交溶液中，于 42℃下杂交 2 h。

杂交后微阵列的清洗与染色

13. 将杂交好的 Tag Array 从杂交箱中取出，按照以下方法清洗：
    a. 用 1×SSPE-T 漂洗阵列两次，每次 10 s。
    b. 在转速为 40～50 r/min 的电动旋转架上，用 1×SSPE-T 于 40℃下清洗芯片 15～20 min。
    c. 22℃下，用 6×SSPE-T 在流体工作站中清洗阵列 10 次。

    也可以用 GenFlex Array 提供的脚本，直接在 GeneChip 流体工作站中清洗阵列。它们的使用方法在 FS400 用户手册中有介绍。
14. 室温下，在转速为 40～50 r/min 的电动旋转架上，用 120 μl 的染色溶液将阵列染色 15 min。
15. 染色后，在 22℃下，用 6×SSPE-T 在流体工作站中漂洗探针阵列 10 次。

427 **微阵列的扫描**

16. 使用定制的，分辨率为每个单元 60～70 像素的扫描系统，在 530 nm 和570 nm 下扫描微阵列。
17. 使用 GeneChip 软件将图像文件转换为数据文件，从而得出一个列表，列表中是微阵列中的每条探针强度的数值。

数据分析

在 GenFlex 阵列上合成的序列是经过仔细选择的，富含在与错配靶分子杂交时具有高的内在分辨能力的探针（见 GenFlex Tech Note 1，Affymetrix 700482，Rev. 2）。但是，在序列复杂性很高的样品中，由浓度造成的与靶分子相似但并不一致的序列的结合而产生的信号偏差仍然有可能发生。因此可以为每条俘获探针（capture probe）设置一条单碱基错配形式作为对照，然后从观测到的整体信号中扣除这种偏差信号。对于完全匹配（PM）和错配（MM）的探针来说，偏差信号的大小通常是相同的。由于 GenFlex 俘获探针的内在的分辨能力是如此之高，以至于扣除很少就会造成超出微不足道的信息损失。因此，每个被俘获的靶分子的原始数据，完全可以被用作每个发射波长（530 nm 和 570 nm）的近似值，在 GeneChip Hybridization Analysis Window 报告中列举为：$S_{570} = PM_{570} - MM_{570}$ 和 $S_{530} = PM_{530} - MM_{530}$。由于藻红蛋白的信号中会有部分与荧光素（“530”）信号通道发生重叠，反之亦然，可以用根据包含在杂交反应中的内标所估计的交叉干扰校正来进行更正。对于这两种标记，单次迭代、通道比率的扣减通常已经足够了。例如，在杂交反应中，可以存在一组 20 条荧光素标记的寡核苷酸（Genset）其与 GenFlex 对照探针（如 GenFlex Technical Note 中所介绍）的一个子阵列互补，浓度为 50 pmol/L。还可以有一组重叠的 20 条生物素标记的寡核苷酸（Genset），其中 10 条与荧光素标记的寡核苷酸具有相同的序列，另外 10 条与 GenFlex 对照探针的一个特定的子阵列互补，浓度为 50 pmol/L。对于每个只有荧光素标记的靶分子杂交的单元，要计算出 S_{570}/S_{530} 强度比值，并对全部 10 个数据进行平均以得到 $F_{570/530}$ 修正因子。相反，对

于每个只有生物素标记的靶分子杂交的单元，可以对10个相应的$S_{530/570}$比值进行平均来计算藻红蛋白修正因子$P_{530/570}$。经过校正的信号数值可以用$C_{530}=S_{530}-S_{570}\times P_{530/570}$以及$C_{570}=S_{570}-C_{530}\times F_{570/530}$来计算。

这些校正算法应该也可以用于与等摩尔的生物素和荧光素标记寡核苷酸混合物杂交的10个单元各自的570 nm和530 nm通道的信号，并且通过对比值C_{570}/C_{530}的平均得到标准化因子$N_{570/530}$。该数值只需要大概的数值，就可以标准化两种光量子由于产率不同的标记所带来的内在的强度差异，并为下面的步骤平衡其大小，用来计算可以反映SBE结果的每个单元的相对等位基因信号（relative allele signal，RAS）：$RAS=N_{570/530}\times C_{530}/(N_{570/530}\times C_{530}+C_{570})$。RAS值本质上反映了在每个单元总的信号中，由荧光素标记的SBE产物所贡献的那部分信号，并且如果产生生物素标记靶分子的等位基因的基因型是纯合子，该值趋于零；如果荧光素标记的互补核苷酸对应的基因型是纯合子，该值趋于1；那么杂合子将会得到中间值。对于特定的位点，RAS值趋于一个典型值，由于受到局部序列特征或者结构对引物结合的影响以及不同核苷酸参入效率的差异，这个数值不一定是精确到0.5。

评价芯片与传统的SNP方法之间的不一致率依赖于SNP位点内在的因素、训练集合的大小、该训练集合现有的基因分型参考的质量、算法的完备程度等。同时，不一致率也依赖于进行比较的方法的准确性。关于该技术精确率更为详细的讨论，见方案8。通常，大多数灵活的基因分型方法的一致率可以表示为大于98% ×大于98%≥96%（见Fan等人的讨论，2000）。要注意到为一种特定的标记集而高度优化的检查一定可以产生较高的准确率。

428

单碱基延伸的疑难问题解答

SBE整体上的性能通常取决于扩增以及SBE反应是否成功。通过预先对位点进行筛选，并且仔细优化引物的设计以避免二级结构，就可以获得高的通过率（90%）及精确度（98～99%）（J. Hirschhorn，私人通信）。要考虑的参数有PCR及SBE引物的熔解温度、长度、自我折叠以及潜在的二聚体形成。在杂合性位点处的引物错配率低，但仍然可以检测到，检测率依赖于含有待分析SNP位点的DNA样品在整体上的复杂程度，其在该位点处可能是复合的。要根据预想的反应多重程度，进行位点特异性扩增和SBE标签寡核苷酸的设计。对于位点特异性扩增，引物的选择要尽量靠近多态性位点以保证多重PCR的成功率最高，理想情况下要扩增出长度小于120 bp的均一的扩增子。通常只需要选择一条链用于SBE反应，但是含有另一条链的独立反应可以建立冗余，并能改善准确性。

荧光基团的选择和相关的光谱重叠也是用户要考虑的关键问题。一旦选定了荧光基团，就要设计一个标记物浓度矩阵，来对发射光谱潜在的重叠进行取样和校正。假定出SNP的双等位基因特性，当用两种荧光基团来区分所有SNP的两个等位基因时，最少有6种反应方式（A/G、A/T、A/C、G/T、G/C和T/C）。向该反应中加入未标记的ddNTP，就可以提高检查的特异性。

其他需要考虑的参数包括为芯片上得到的信号指派具有特定置信水平的基因型读出（genotype call）时所需要的分析与聚类算法。对于一个足够大的样本集合，任意一个位点的RAS分值应该聚类到最多三个集合中的其中之一，这三种集合是对应于一个双等位基因标记可能的三种二倍体基因型。如果聚类的区分足够明显，就能够为独立地分析样品各组成成分的基因型提供足够的信息，同时也提供了划定基因型读出边界的训练，并可以用于后续的样品。在某些情况下，对于一种特定的标记，借助于从两条链上的SBE得到的RAS评分进行二维作图，就可以获得额外的聚类分辨能力。

429

方案 8

高密度微阵列用于序列变异检测

Thomas Ryder，Hajime Matsuzaki，Suzanne Dee（昂飞公司，圣塔克拉拉，加利福尼亚，95051）

用短链寡核苷酸阵列来检测基因或基因组的连续核苷酸的方法已经发展成为一种高效的检测方法（Chee et al. 1996；Kozal et al. 1996）。这种策略，通常称为再测序（re-sequencing），需要产生一套在基因上呈叠瓦式排列的寡核苷酸阵列，长度通常为 15～25 个碱基，共同地覆盖于一个已知基因的全部序列上。为了确定任一给定碱基的序列，就需要生成四条与感兴趣的基因区域互补，但是彼此之间在靠近寡核苷酸的中间位置上各不相同的寡核苷酸探针，在这些探针中该位置被所有四种可能的碱基（A、G、C、T）分别替换。四个寡核苷酸通常称作一个探针组，并且在一个典型的阵列中包含了数千个这样的探针组。从感兴趣的基因得到的荧光标记的靶分子通常与每组变化中正确匹配的探针杂交最强，可以据此“读出”该位置上的序列。读出阵列中其他单元上的全部杂交模式，可以使得靶 DNA 的全部序列得到确定。通过应用复杂的统计学和模式识别算法，以及多套不同长度和（或）取代位置的探针组进行目标碱基的冗余检测，可以增加精确度和可信度。如果在阵列上生成可以覆盖突变导致的新序列的单元，那么这种方法也适用于检测基因中已经很清楚的缺失或者特定的插入突变（图 6-10）。

在检测基因组中特定区域上新的自发突变或者是辨别可遗传的单核苷酸多态性时，再测序技术是一种有力快速的方法（Hacia et al. 1996；Wang et al. 1998；Ahrendt et al. 1999；Hacia and Collins 1999；Wikman et al. 2000）。使用现有的光刻合成技术，并结合长链 PCR 产物的扩增，在单次杂交反应中有可能对长度从 10 kb（Cutler et al. 2001）到 10 000 kb（Patil et al. 2001）的区域进行部分的再测序；后一个规模依赖于将整块晶片与杂交样品一起孵育。

本方案中介绍了一种分析人 *p*53 基因全部编码序列，来检测自发突变的方法，这种方法即使是在存在中等数量的来自正常组织中野生型序列污染的情况之下也可以有效检测。该方案说明了应用寡核苷酸阵列可以同时分析数百万碱基的原理，并且可以用便利的商品化试剂来进行实验。

*p*53 突变扫描方法回顾

可以用各种实验室标准技术从实验样本中提取 DNA。作者的实验室中使用 QIAGEN 的 QIAamp 组织试剂盒，得到了可靠的结果。要注意来自甲醛固定组织中的 DNA 很可能受到破坏，需要以下介绍的方法进行特殊处理（见 Greer et al. 1993；Greer et al. 1995）。用抽提的 DNA 作为模板，*p*53 基因全部 10 个外显子在一次反应中都被扩增，然后这些扩增子被片段化并在 3′端标记上带有荧光素的双脱氧核苷酸。用杂交缓冲液稀释标记的片段，缓冲液中含有一个荧光素标记的寡核苷酸，与分布在整个阵列上的探针互补，因此用作杂交的对照。

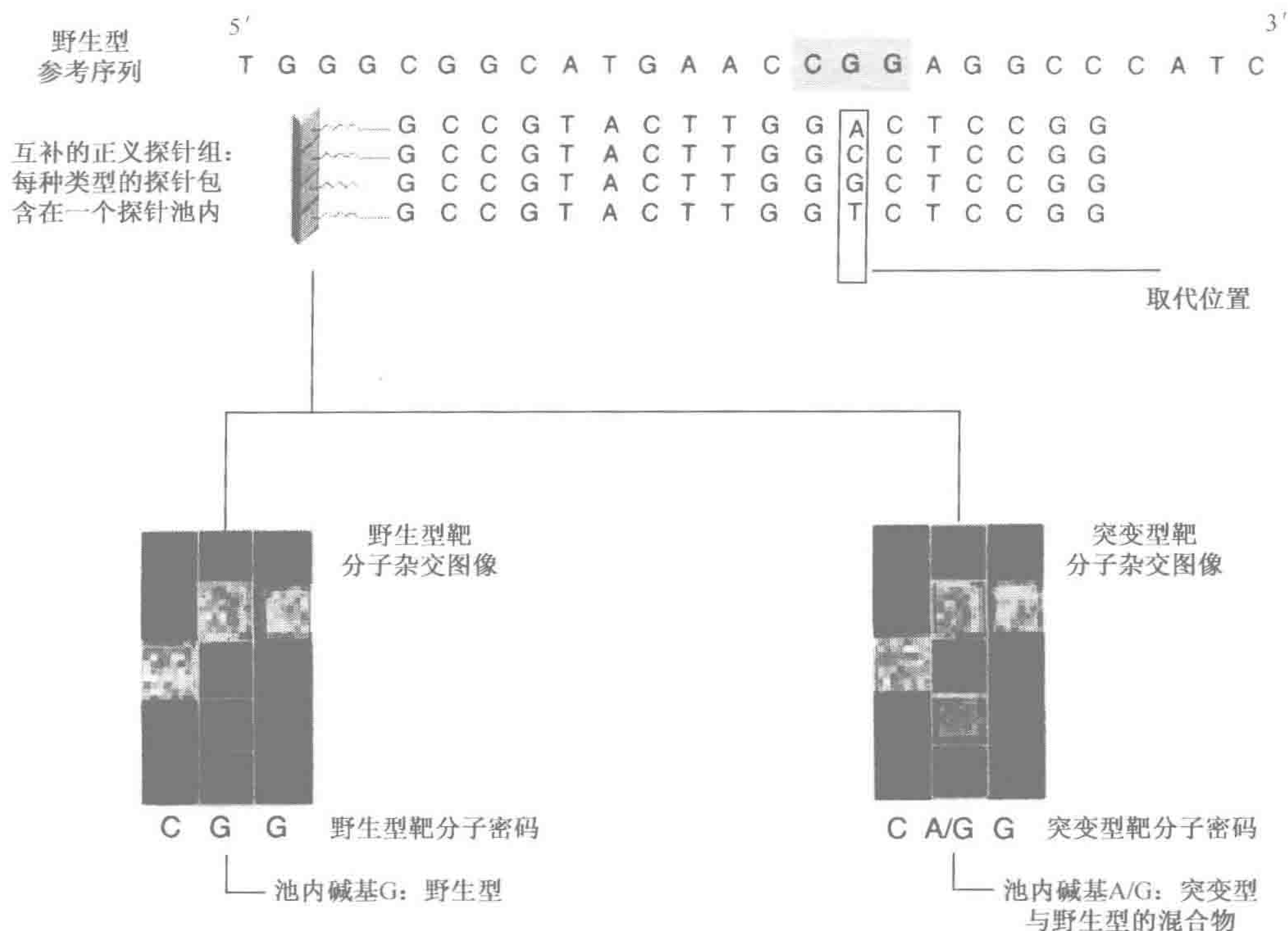


图 6-10　芯片阵列系统示意图

阵列中的探针按照一组四个进行排列。每组中的每条探针除了一个错配位置以外，与参考序列互补，该位置称作“取代”位置。在此位置上，四种可能的核苷酸（A、C、G、T）在探针组中得到反映。分析条件优化了荧光标记的 DNA 靶分子和与其序列最佳互补的探针之间的杂交。这个正确匹配的杂交体相对于组内其他三个（译者注：原书误为四个）靶分子-探针杂交体会产生更高的荧光强度。芯片具有与 *p*53 基因中每个碱基互补的基因［根据 Ahrendt et al. 1999 修改（经 Affymetrix 允许）］

430

通过低复杂度的 PCR，可以得到足量的靶分子，就可以进行只有 30～60 min 的杂交反应。通过自动化的流体工作站控制样品以及清洗缓冲液的导入、排出和混合，以及孵育温度，这样就使样品处理变得容易，标记的 DNA 与阵列上的探针在缓冲溶液中杂交30 min，然后清洗。

431

杂交后的探针阵列用 GeneArray Scanner 扫描仪进行扫描。扫描耗时约 4min，产生一个数据图像文件（*.dat）。用 Affymetrix Microarray Suite 软件对图像进行数据分析。在 Image（图像）窗口中，网格自动地加到图像上，并通过软件计算出每一探针池的强度。将从样品序列得到的杂交强度与从参考序列得到的结果进行比较。鉴别出强度模式不同于参考标准的，显示包含有突变碱基的位点。有关使用流体工作站、扫描仪和分析软件更多的背景信息可以从特定的用户指南和在线帮助（www. affymetrix. com）中得到。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

重要事项：配制以下试剂时使用分子生物学级（无核酸酶）的 H_2O。

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

乙酰化牛血清白蛋白（BSA）（20 mg/ml）（B 8894，Sigma）；

EDTA（0.5 mol/L）（15575-012，GIBCO）

无水乙醇

H_2O，分子生物学级（16-001Y，BioWhittaker）

浓缩杂交液（12×SSPE/0.1% Triton X-100）

将 600 ml 20×SSPE 与 390 ml H_2O 完全混合，加入 10 ml 10% Triton X-100，充分混合后用 0.2 μm 或 0.45 μm 的真空滤器过滤。将体积为 1 L 的缓冲液按照每管 50 ml 的量分装到 20 只聚丙烯离心管中，室温下保存。

$MgCl_2$（25 mmol/L）（N808-0241，Applied Biosystems）

4×PCR 试剂

10×GeneAmp PCR 缓冲液Ⅱ	5 ml
25 mmol/L $MgCl_2$	5 ml
100 mmol/L dNTP	100 μl

加入 2.1 ml 的 H_2O 使总体积达到 12.5 ml。充分混合溶液，向 50 个微型离心管中按每管 250 μl 的体积分装，－20℃下保存。

重要事项：在实验室中指定的无模板区域制备和保存 4×PCR 试剂。

20×SSPE（16-010Y，BtoWhittaker）

Tris-乙酸盐（400 mmol/L，pH 8.2）（0279-IL-E，Teknova）

Triton X-100（T 9284，Sigma）

清洗缓冲液 A（3×SSPE/0.005% Triton X-100）

向 850 ml 的 H_2O 中加入 150 ml 的 20×SSPE 和 500μl 10% Triton X-100，充分混合并用 0.2 μm 或 0.45 μm 的真空滤器过滤。室温下保存。

酶和酶缓冲液

AmpliTaq Gold（N808-0241，Applied Biosystems）

本方案是为使用 AmpliTaq Gold 的反应进行的优化，建议使用这种酶。

BioArray Terminal Labeling w/FL ddCTP 试剂盒（Enzo，由 Affymetrix 销售，900180），用来标记片段化的 DNA 扩增子。

小牛肠碱性磷酸酶（18009-027，GIBCO）

GeneAmp 10×PCR 缓冲液Ⅱ（N808-0241，Applied Biosystems）

10 mmol/L Tris-HCl（pH 8.3）

50 mmol/L KCl

片段化试剂（随 GeneChip p53 试剂盒提供，900132）

在 10 mmol/L Tris-HCl（pH 7.5），10 mmol/L $CaCl_2$，10 mmol/L $MgCl_2$ 和 50%甘油中配制 DNase Ⅰ酶溶液。

核酸和寡核苷酸

DNA 分子质量标记（50 bp）（10416-014，GIBCO）

dNTP 储存液（每种 100 mmol/L）（27-2085-1，Pharmacia Biotech）

GeneChip p53 试剂盒（900132，Affymetrix）

试剂盒组分：

- p53 引物集合［20 种引物，溶于 1 mmol/L Tris-HCl（pH 7.4），0.1 mmol/L EDTA 中］

  Exon 2：5′-TCATGCTGGATCCCCACTTTTCCTCTTG-3′
  　　　　5′-TGGCCTGCCCTTCCAATGGATCCACTCA-3′
  Exon 3：5′-AATTCATGGGACTGACTTTCTGCTCTTGTC-3′
  　　　　5′-TCCAGGTCCCAGCCCAACCCTTGTCC-3′
  Exon 4：5′- GTCCTCTGACTGCTCTTTTCACCCATCTAC-3′
  　　　　5′-GGGATACGGCCAGGCATTGAAGTCTC-3′
  Exon 5：5′- CTTGTGCCCTGACTTTCAACTCTGTCTC-3′
  　　　　5′-TGGGCAACCAGCCCTGTCGTCTCTCCA-3′
  Exon 6：5′-CCAGGCCTCTGATTCCTCACTGATTGCTC-3′
  　　　　5′-GCCACTGACAACCACCCTTAACCCCTC-3′
  Exon 7：5-GCCTCATCTTGGGCCTGTGTTATCTCC-3′
  　　　　5′-GGCCAGTGTGCAGGGTGGCAAGTGGCTC-3′
  Exon 8：5′-GTAGGACCTGATTTCCTTACTGCCTCTTGC-3′
  　　　　5′-ATAACTGCACCCTTGGTCTCCTCCACCGC-3′
  Exon 9：5′-CACTTTTATCACCTTTCCTTGCCTCTTTCC-3′
  　　　　5′-AACTTTCCACTTGATAAGAGGTCCCAAGAC-3′
  Exon 10：5′-ACTTACTTCTCCCCCTCCTCTGTTGCTGC-3′
  　　　　 5′-ATGGAATCCTATGGCTTTCCAACCTAGGAAG-3′
  Exon 11：5′-CATCTCTCCTCCCTGCTTCTGTCTCCTAC-3′
  　　　　 5′- CTGACGCACACCTATTGCAAGCAAGGGTTC-3′

- 对照寡核苷酸 F1（溶解于 10 mmol/L Tris-HCl，pH 7.4，1 mmol/L EDTA 中，浓度 100 nmol/L）序列 5′-荧光素-CTGAACGGTAGCATCTTGAC-3′；
- 片段化试剂

  溶解于 10 mmol/L Tris-HCl（pH 7.5），10 mmol/L $CaCl_2$，10 mmol/L $MgCl_2$，和 50%甘油中的 DNase Ⅰ
- p53 参照 DNA

  人胎盘 DNA（50 ng/μl），溶解于 10 mmol/L Tris-HCl（pH 8.0），0.1 mmol/L EDTA（pH 8.0）中
- p53 探针阵列

  我们建议对于参照样品要用同一批次的探针阵列，并且在每个实验中实验样品要和一个参照样品进行比较。

来自组织样本的基因组 DNA，使用如 QIAamp 组织试剂盒（29304，QIAGEN）进行分离。

将分离的 DNA 用 10 mmol/L Tris（pH 9）稀释为浓度 5～50 ng/μl，4℃下保存。

433 **凝胶**

微型琼脂糖凝胶电泳装置，含有合适的缓冲液和溴化乙啶〈!〉
NuSieve 琼脂糖（3∶1）（50092，FMC Corporation）

专用设备

离心管（锥形，50 ml）（25325-50，Corning）
Eppendorf 微量离心管（1.5 ml）（20901-641，VWR）
GeneArray 扫描仪（Affymetrix）
GeneChip 流体工作站 400
MicroAmp（0.2 μl）PCR 管，带管盖（N801-0612，Applied Biosystems）
Microarray Suite Analysis 软件（Affymetrix）
微量移液器（P-2、P-20、P-200、P-1000）（Rainin Pipetman 或者与之相当的）
移液器吸头（带有灭菌滤芯）
热循环仪和附件（Applied Biosystems GeneAmp PCR System 2400 或 9600）
真空过滤装置（容量为 1 L，滤孔 0.20 μm 或 0.45 μm）（25988-1L，Corning）

方法

重要事项：将扩增方案的步骤分开，并在实验室中三个独立的区域中进行，以减少将前次扩增反应的 PCR 产物带入的概率。

为抽提的基因组 DNA 建立扩增：无模板区域

在指定的无模板区域进行最初的 PCR 预混试剂和反应试管的制备。使用专用于该操作的移液器、吸头和试管架。将 4×PCR 试剂、GeneChip p53 引物集和 *Taq* 聚合酶保存于指定的无模板的－20℃冰箱里。

1. 按照以下比例将 PCR 预混试剂各组分混合。每个反应需要 50 μl 的 PCR 预混试剂。反应的总数等于组织样品和参考的数量，再加上两个额外的反应：一个是对照，另一个是用于补充加样时的损耗。

组　分	体积/μl
4× PCR 试剂	25
p53 引物集	5
AmpliTaq Gold（5 单位/μl）	2
H_2O，分子生物学级	18
PCR 预混试剂体积	50

让各组分完全解冻。吸取 Ampli*Taq* 聚合酶时要格外小心，因为它是储存在 50%的甘油中。

2. 向 PCR 管中各加入 50 μl PCR 预混试剂。盖紧管盖。

准备扩增反应：少量模板区

434

3. 按以下步骤向装有 PCR 预混试剂的试管中加入样品和 GeneChip p53 参考 DNA。
   a. 向反应管中的 50 μl PCR 预混试剂中加入 50 μl H_2O，用于无模板的对照反应。盖紧管盖。
   b. 向反应管中的 50 μl PCR 预混试剂中加入 5 μl GeneChip p53 参考 DNA 和 45 μl H_2O，用于参照 DNA 反应。盖紧管盖。
   c. 在 50 μl 体积中分别加入 250 ng 实验样品 DNA 和水，用于实验样品。盖紧管盖。每个 PCR 的总体积为 100 μl。

扩增和产物分析：大量模板区

将用于 PCR 的热循环仪和电泳设备放在实验室中为大量模板的指定区域。

4. 将试管放入热循环仪中，按照下面的变性、复性、聚合时间与温度进行扩增反应：

循环次数	变 性	复 性	延 伸
1	96℃，10 min		
35	95℃，30 s	60℃，30 s	72℃，45 s
1			72℃，10 min

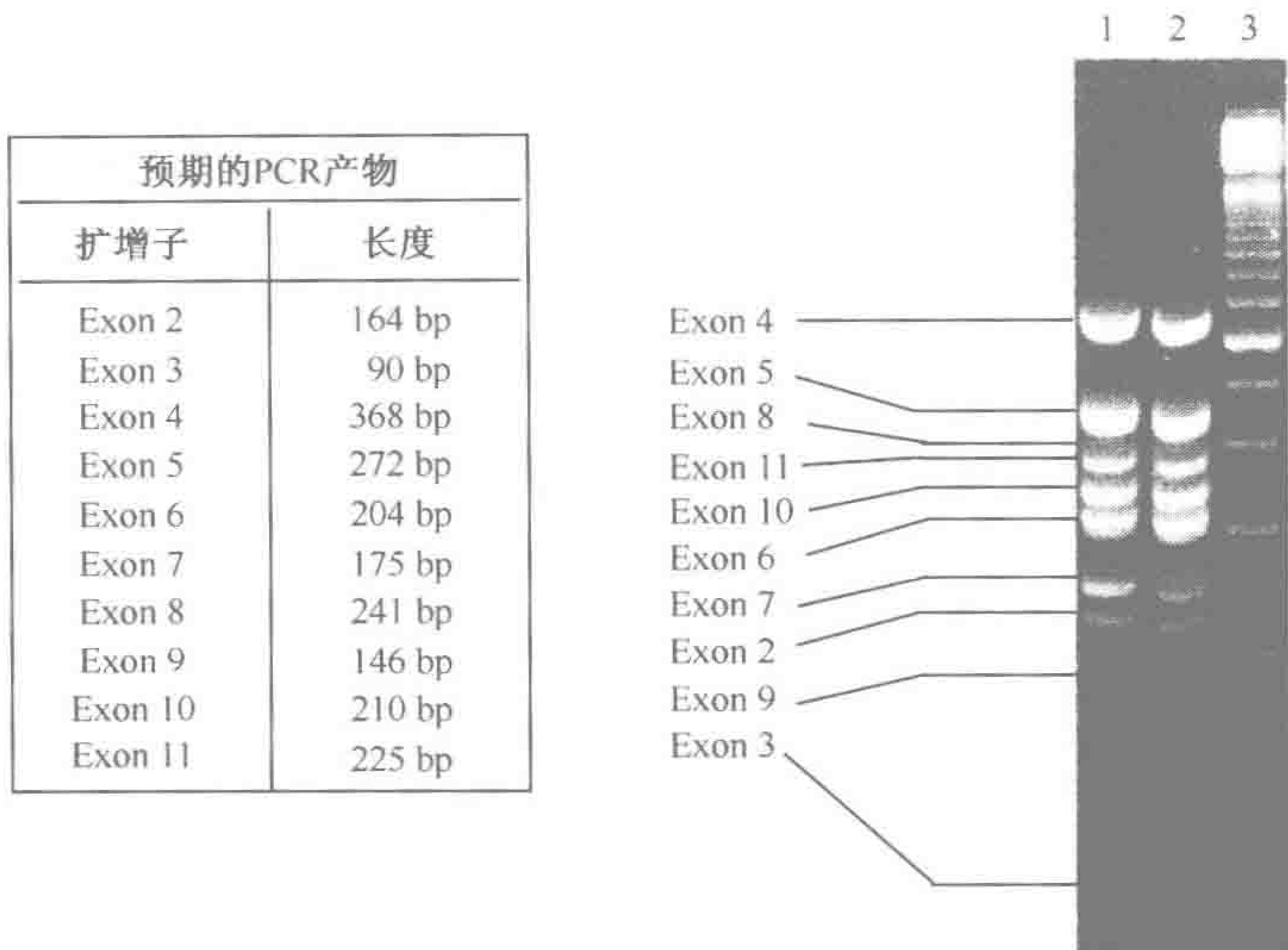
将反应从热循环仪中取出，4℃下保存直到用于以下步骤。

这些时间和温度是为本方案中特定的设备和反应而优化的（Applied Biosystems GeneAmp PCR System 2400 或 9600 热循环仪），需要对条件进行修改以适合其他类型的设备和反应。

5. 4% NuSieve 琼脂糖凝胶电泳检测 PCR 产物，加入 50bp 的 DNA 分子质量标记。
   a. 用至少 5μl PCR 产物与合适的凝胶上样缓冲液进行上样，使用 1μl 50bp 的 DNA 分子质量标记和凝胶上样缓冲液上样。
   b. 在 1×TBE 中，于 125 V 下电泳 60～90 min。
   c. 将凝胶在溴化乙啶（0.5 μg/ml，溶解在 1×TBE 中）中染色，观察 DNA 的条带。

   如果在上样时每种 PCR 产物不超过 5 μl，不一定所有的条带都能看见。

6. 在继续进行下面的步骤之前，检查凝胶以确保 PCR 产物的质量和产量（图 6-10，图 6-11）。关于各种问题处理的建议，见方案最后疑难问题解答的信息栏。



435

图 6-11　参考（第一泳道）DNA 或实验（第二泳道）DNA 的 *p53* 基因编码外显子经过 PCR 扩增后的片段的凝胶电泳和溴化乙啶染色结果图

DNA 分子质量标记（第三泳道）用于估测片段的大小和扩增子的丰度。要注意，得到的是分离的片段，而不是扩散的模糊条带。一般从高质量的模板可以得到 1～3 μg 的 DNA 片段。表格中给出了预期的 PCR 产物

DNA 扩增子的片段化

GeneChip 片段化试剂的最合适的用量为 0.15 单位/μg DNA。下面的片段化反应适合于 1～3 μg 的 DNA，正好是从高质量的模板得到的 PCR 产物的产量。关于处理诸如一致的弱信号强度或者分辨较差等问题，见本方案最后疑难问题解答的信息栏。

7. 将 45 μl 样品 PCR 和参考 DNA PCR 产物转移到一支新的 MicroAmp 反应管内。为了得到更为精确的结果，将另外的 45 μl 各 PCR 产物转移到另一组反应管中以进行重复分析。否则，将剩余的 50 μl 各 PCR 产物在 4℃下保存，以备后用。
8. 按照下表配制片段化反应的混合物：

组　分	足够供给 24 个反应的混合物	终浓度
片段化试剂（检查产品标记上给出的浓度）	6 单位	0.25 单位
EDTA/（20 mmol/L）*	x μl	可变的
碱性磷酸酶/（1 单位/μl）	60 μl	2.5 单位
Tris-乙酸盐/（10 mmol/L）	从 60 μl 中减去所加入的片段化试剂的体积	0.5 mmol/L
工作混合物体积	120 μl	

*使用 Affymetrix 片段化试剂时，EDTA 的用量要与酶的用量成比例，以复合酶保存缓冲液中的过量二价离子，Affymetrix 从供应商获得 DNase Ⅰ。

反应的数目等于样品和参考的数量加上一个多余的反应，以补偿操作时的损耗。移取片段化试剂和磷酸酶时要特别仔细，因为这些酶保存在甘油之中。

重要事项：工作混合物必须现配现用，用后丢弃剩余的液体。

9. 向45 μl PCR 产物中加入 5 μl 工作混合物，至反应体积为 50 μl。

   加入工作混合物要迅速，以减少第一管反应和最后一管反应在孵育时间上的差异。建议使用多段分注加样器或者多通道加样器。将样品充分混合。

10. 将反应管放入热循环仪中，按照下列的时间和温度编程：

循环次数	时　间	温　度
1	15 min	25 ℃
1	10 min	95 ℃

将反应管从热循环仪中取出，置于 4℃下保存直到进行末端标记反应。如果要长期保存扩增子，将反应物在－20℃下保存。

该反应在 Applied Biosystems GeneAmp PCR System 2400 或 9600 热循环仪上进行了优化。因为该反应的孵育时间和热灭活非常关键，建议不要使用简单的加热槽。

标记片段化产物

片段化的 PCR 产物可直接用作荧光素标记（或其他的荧光基团或者半抗原）核苷酸的掺入反应的底物。

11. 用末端转移酶和 Enzo BioArray Labeling 试剂盒中提供的组分在片段化的扩增子的 3′端标记上带有荧光素的双脱氧核苷酸。

本步骤中建议使用 Enzo BioArray Labeling w/FL ddCTP 试剂盒，该试剂盒内提供了酶和一种标记的核苷酸。请参考试剂盒内提供的方案进行标记反应。

样品与 GeneChip p53 阵列的杂交

GeneChip p53 检查需要用一个探针阵列对 DNA 样品的正义链和反义链进行分析。GeneChip 检查的试剂是为在 GeneChip Fludics Station 400 和 GeneArray Scanner 中使用而设计的。

12. 按照下面的比例混合杂交预混试剂。杂交反应物的数目等于样品和参考的数目，再加上一份额外的试剂用来补偿移液误差。

组　分	体积/样本	杂交中的最终浓度或数量
杂交缓冲液浓度	250 μl	6×SSPE
乙酰化 BSA/（20 mg/ml）	50 μl	2 mg/ml
对照寡核苷酸 F1/（100 nmol/L）	10 μl	2 nmol/L
H_2O	90 μl	—
杂交预混试剂最终体积	400μl	—

13. 将杂交预混试剂按照每份 400 μl 分装到微量离心管中。

437

14. 将 100 μl 的标记产物（步骤 11*）加入装有 400 μl 杂交预混试剂的微量离心管中，使最终杂交体积为 500 μl。充分混匀试剂。

杂交的靶分子样品可以在－20℃下最多避光保存 4 周。避免重复冻融。

靶分子杂交和清洗

15. 建立杂交实验和流体工作站

    GeneChip Fluidics Station 400 由 Affymetrix Microarray Suite 软件控制，可以在探针阵列上进行杂交、染色和清洗程序。每个 GeneChip 检查都有相应的仪器控制脚本，指导每个芯片反应室（cartridge）内杂交和清洗溶液的注入和搅拌，以及这些步骤的时间和温度控制。有关调整检查的条件的更多的细节和介绍在随同 GeneChip p53 阵列提供的技术资料中有介绍。

16. 在 45℃下，将含有荧光标记 DNA 片段的溶液与 p53 探针阵列杂交 30 min。按照 Affymetrix GeneChip *p53* 手册中的说明进行样品与阵列的杂交，以及阵列的清洗。
17. 在清洗循环开始之前，用清洗缓冲液 A 冲洗 GeneChip Fluidics Station 400 的管道。
18. 清洗时，在 35℃下用清洗液 A 按每次循环 10 个混合（mix）进行两次循环。然后保持于 20℃。

    在杂交和清洗程序完成之后，探针阵列可以直接进行扫描。

19. 用 GeneArray Scanner 扫描杂交后的阵列。

    扫描仪用 Affymetrix Microarray Suite 软件控制。关于扫描更多的信息，参照 Affymetrix Microarray Suite 用户指南和在线帮助，以及 GeneArray Scanne 用户指南。探针阵列扫描结束后，将得到的图像数据用（.DAT）格式文件命名保存在 GeneChip 工作站的硬盘上。

样品分析

数据分析是建立在将未知样品的杂交模式与野生型参考进行比较的基础之上的。用探针池中包含的每个像素点的强度，通过软件计算出每个探针池的平均强度。每个探针池边缘上的像素没有包含在内，这样可以避免邻近的探针池对从单个探针池计算出的平均强度的影响。给计算得到的平均强度指派一个 x/y 坐标位置，与探针池在阵列中的位置相对应。这些数据以和（.EXP）与（.DAT）文件相同的名字保存为（.CEL）文件。（.CEL）文件是一种中间数据文件，在 Affymetrix Microarray Suite 中不显示。

然后在软件中应用所选择的探针阵列算法对探针池的强度进行比较，并确定每个位置上的碱基。这个过程要参考探针阵列的第二个库文件，.CDF 文件内的信息来完成。在 Affymetrix Microarray Suit 的序列分析窗口（Sequence Analysis window）内，样本中核苷酸变化的总结报告以（.CHP）文件形式自动显示。.CHP 文件的名称与（.EXP）、（.DAT）和（.CEL）文件相同。

438

疑难问题解答和特别注意事项

- **扩增产物质量的评价。**原理上，大的缺失或其他重排在凝胶条带谱上显示得很明显，但是这类改变在 Soussi p53 突变数据库中（http://p53.curie.fr）很少出现。因此，电泳图谱的改变更可

* 原书误为步骤 12——译者注

能是反映了与 PCR 有关的问题（图 6-10 和图 6-11）。这类问题最有可能反映了初始基因组 DNA 的质量问题。当使用抽提自甲醛固定和石蜡包埋组织中的基因组 DNA 时，要获取足够量的长度大于～200 bp 的扩增产物是有问题的。这会导致在凝胶上发生条带的缺失或者一些条带强度降低。如果出现这种情况，可能需要重新进行扩增。

- **DNA 扩增子的片段化。**在有小牛肠碱性磷酸酶存在的情况下，用 DNase Ⅰ消化 DNA 扩增子进行片段化。片段化可用来促进靶分子与探针的接触，缩短所需要的杂交时间，以及改善对完全匹配/错配信号的识别。片段化的程度对基于芯片再测序的准确性非常关键。因此要通过多因素滴定实验来建立该实验的最佳条件，以使得错配识别性能和整体上的准确性达到最高（见 Cutler et al. 2001）。经过实验确定，GeneChip 片段化试剂的最佳用量为 0.15 单位/μg DNA。片段化反应适合于消化 1～3 μg 的 DNA，该数量相当于通常从高质量模板扩增得到的 PCR 产物的量。如果杂交后观察到一致性的弱信号强度或者识别效果不好（特别是参考 DNA），应该测定 PCR 反应中 DNA 的量，并根据每微克 DNA 使用 0.15 单位片段化试剂来重新计算合适的片段化反应成分。
- **数据分析：检查的特异性。**通常，再测序应用需要很高的准确度以确保任何候选变异的预测值都能被报告到。一般的，每个核苷酸的假阳性率×所用核苷酸数目表示的再测序靶分子的长度必须要远小于预期的变异率。例如，即使是名义上的 99.9%的高特异率，对长度为 1.5 kb 的序列进行再测序得到的报告里，分析的 *p*53 基因中 80%有一个或者更多的突变，该结果远远高于真正的自发突变率。

GeneChip p53 分析能够提供有效扩增的 10 重 PCR 和优化的杂交条件，可以从 *p53* 基因中富含 A/T 区域得到可靠的信号，同时检查大多数富含 G/C 基因区域所用的探针都具有较高的鉴别能力，并且相互之间没有交叉杂交的假象。通过冗余检查，准确性也得到提高。全部位置至少检查两次，在每条 DNA 链上检查一次。在最关键的位点上，每条链有 7 个额外的探针组，相对于探针末端使用了不同的取代位置，这主要是用于 300 多个最普遍的 *p53* 突变的检测（根据 Soussi et al. 的编辑；见http://p53.curie.fr）。最后，突变检测不是根据假定的在野生型和突变分析的五种探针（也就是说，T、G、C、A 和所检查的碱基缺失）之中的强度分布，而是通过将它们和包含在每批样品检查当中的野生型参考样品进行比较来实现。相对于从参考样品得到的杂交模式，实验样品中的强度模式发生了与经验所确定的强度模式相比一定程度的改变时，就表明存在突变。如同这里所述，参考标准可以用来比较两个不同检查的杂交模式，或者，在单次杂交中，也可以将分别带有不同的、可区分标记的实验和标准样品混合在一起来进行实验。

在 Affymetrix 的研究中，对样品进行一次实验，GeneChip p53 检查总体上的准确性约为 99.5%（与双脱氧测序法比较）。因为假阳性是随机发生的，对样品进行两次分析，并要求两次分析具有一致的突变读出（mutation call），以此来界定一次检测事件，这样可以将准确性提高到大于 99.94%，并且特异性大于 99.99%。检测突变的灵敏度在有冗余检查的位点处最为有效。在这些位点处，当混合样品中人为掺入的野生型样品水平小于 25%，通常为小于 15%时，突变可以被可靠地检测到。在一个盲实验中，对来自肿瘤组织的一组性质很清楚的真实患者样品进行突变检测，总体上的灵敏度为 89%。这些结果与双脱氧测序结果的灵敏度很符合（见 Ahrendt et al. 1999；Wickman et al. 2000）。

使用再测序进行单碱基缺失检测时，向每组叠瓦式排列的探针中加入第五条探针，其中的待测碱基被删除，这种方法是很直接明了的，正如在 GenChip p53 阵列中所做的那样。在可以用于训练这个算法的样品集中，很少在样品中观察到这些突变，因此，在该检测的发展过程中，这一潜在的功能没有被充分挖掘。同样，在探针组中增加四条额外的探针，用于检测和当前探针组检查位点相邻的位置上的全部单碱基插入，这也是直截了当的。如果有任何理由可以预期某

个特定的发生缺失的靶分子的大小，多个碱基缺失同样可以通过系统的排布叠瓦式探针来检测。但是，通过系统的排布叠瓦式探针来检测多个碱基插入会带来探针数量呈现指数式增长这一挑战。因此，当分析纯的靶分子时，对于这种类型的复杂突变，基于电泳的测序方法要比与特定的探针杂交更加灵活，但是对于未处理的混合物，如肿瘤样品或者甚至是含有不同等位基因的来自不同病灶的分散二倍体混合物，带有插入或者缺失的样品的序列分析结果也可能无法解释。

439

方案 9

SNP 阵列用于杂合性缺失分析

Kerstin Lindblad-Toh（怀特黑德研究所/MIT 基因组研究中心，剑桥市，马塞诸塞州，02139）

人类癌症的发生是由散发的突变与染色体整体上的变异共同导致的。Knudson 的“二次打击”假说适用于多种类型肿瘤的发生，其中的第一次打击包含了肿瘤抑制基因一个拷贝散发的生殖细胞突变。第二次打击包含基因第二个拷贝的缺失，有时候是通过包含有正常拷贝的染色体发生大片段缺失而发生。随着肿瘤的发展，更多的染色体整体上的变异可能会随之发生。当染色体物质发生大段的缺失时，可以看成是杂合性缺失（loss of heterozygosity，LOH）（Dracopoli and Fogh 1983）。LOH 也可以通过有丝分裂重组、染色体单体断裂诱导的复制、有丝分裂不分离、基因转变或者是很不确定的染色体间重组或缺失的染色体结构变化而发生（Thiagalingam et al. 2001）。

遗传标记，如限制性片段长度多态性（restriction-fragment-length polymorphisms，RFLP）和简单序列长度多态性（simple-sequence-length polymorphisms，SSLP），通常在正常组织的 DNA 中呈杂合性，而在肿瘤组织的 DNA 中呈纯合性。运用 SSLP 的 PCR 扩增和凝胶电泳来进行等位基因分型是在多数不同类型的肿瘤中进行 LOH 分析的一种可靠的方法（Weissenbach et al. 1992）。其他方法，如比较基因组杂交（CGH）（Kallionemi et al. 1994）和光谱核型分析（spectral karyotyping，SKY）（Schrock et al. 1996），涉及对全部染色体进行荧光染色，而且是用来系统的筛查整个基因组中拷贝数的变化。利用遗传标记来研究 LOH 的一个优点在于可以确定基因组中散在位置上的等位基因状态，而 CGH 和 SKY 仅能提供局部的信息。利用遗传标记来研究覆盖整个基因组的大范围的等位基因分型研究需要一组密集的遗传标记物，以及由此而需要的高的基因分型效率。如果可以达到标记密度和基因分型效率的要求，可以通过 LOH 检测微小缺失，并能用于快速鉴别在肿瘤发生中所涉及的基因。

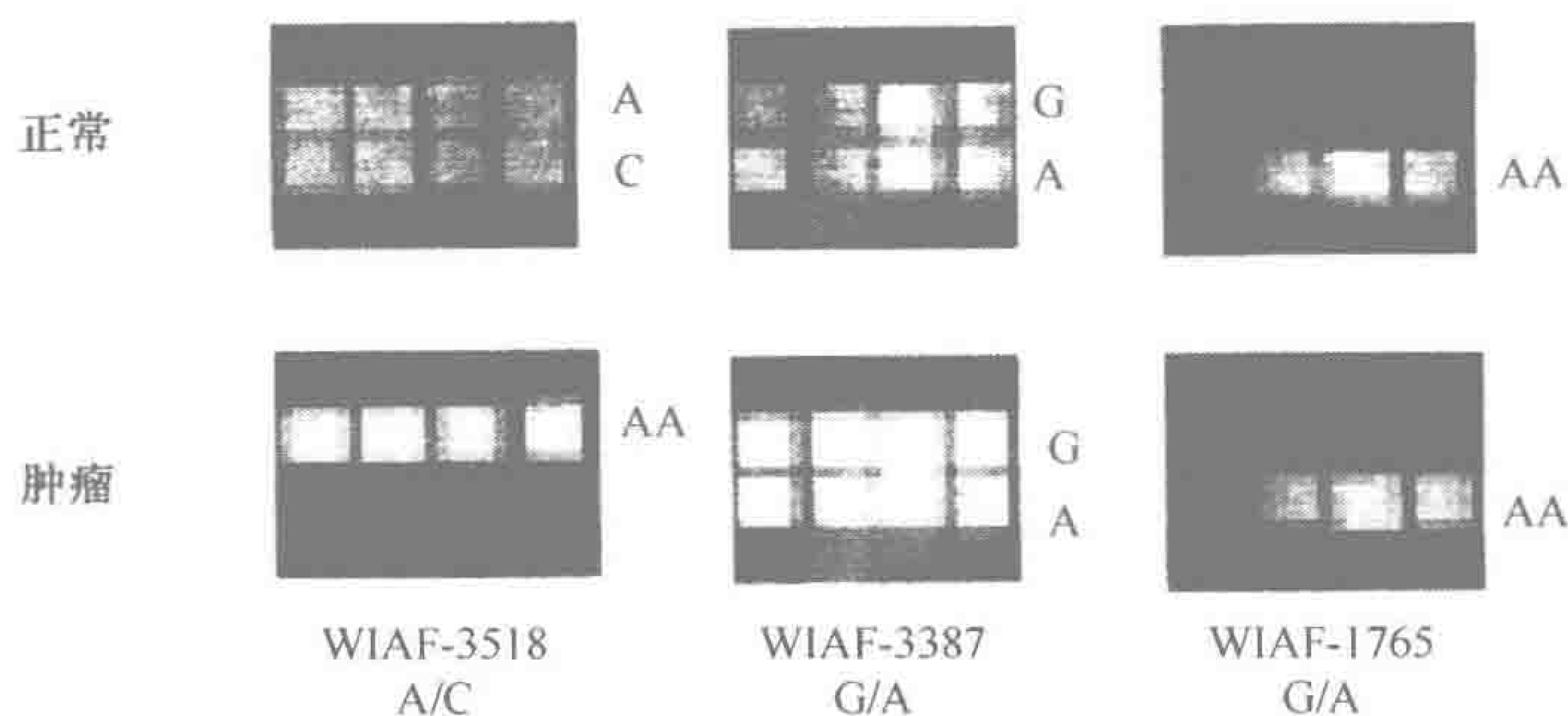
最近在人和其他基因组中发现的大量的单核苷酸多态性（SNP）为获得高密度覆盖整个基因组的标记提供了巨大的贡献。例如，在公共数据库中人类 SNP 的数量已经超过了 160 万，即大约每 2 kb 存在一个 SNP（Sachidanandam et al. 2001）。得益于 SNP 标记的优越性，发展了多种大规模的基因分型方法（关于综述，见 PART Ⅱ 导言中的

表 6-4)。在理论上，许多基因分型技术可以应用于 LOH 分析。但是实际上，所用的方法必须能够克服杂合性的正常组织 DNA 对肿瘤 DNA 的轻微污染。由于其具有灵活性和进行大规模基因分型的能力，质谱（Haff and Smirnov 1997）与基于微阵列的方法很可能在将来成为大规模 LOH 分析的主要技术。

利用相对少量 SNP（600～1500 个 SNP）的基于微阵列的 LOH 分析方法包含两种类型相似的微阵列（Lindblad-Toh et al. 2000a；Mei et al. 2000）。本方案中介绍了其中的一种方法，即与 Affymetrix HuSNP 阵列进行等位基因特异性杂交来检测 LOH（正如 Lindblad-Toh et al. 2000a）。Affymetrix HuSNP 阵列中包含有 1494 个 SNP，是在最初的 SNP 大规模研究项目之一当中鉴定出来的（Wang et al. 1998）。在这 1494 个 SNP 中，约 1205 个可以从不同的人 DNA 样品中得到很好的常规扩增，而且得到的扩增位点所占的比例称为读出率（call rate）。在任何给定的 DNA 当中，大约 30％的 SNP 是杂合性的，因此约有 350 个 SNP 被称为是带有信息的（informative）。虽然在本书写作时，HuSNP 阵列中所包含标记的数量并不是很多，但是使用这些阵列可以对该方法进行原理性验证。此外，在 HuSNP 阵列中达到的标记密度类似于或者高于应用一个常用的杂合度约为 70％的 350 个 SSLP 标记集合时所达到的密度（Cooperative Human Linkage Center [CHLC] Screening set/version 6；CHLC http://gai.nci.nih.gov/CHLC）。

针对个体在 HuSNP 阵列上进行 LOH 时，要对来自正常组织的 DNA（100 ng）进行一次分析，而对肿瘤 DNA（200 ng）进行重复分析。对每份 DNA 样品进行 24 多重 PCR，每份样品中包含了代表 50～100 个位点的位点特异性的 T3/T7 加尾引物对混合物。然后使用生物素化的 T3/T7 引物在 24 多重反应中扩增该反应产物。得到的 PCR 产物合并在一起，与 SNP 阵列杂交，用链亲和素-藻红蛋白染色，并在扫描仪上通过荧光检测进行分析。每个 SNP 位点的检测含有四行 25 聚的寡核苷酸，其中两行含有与 SNP 等位基因 A 或 SNP 等位基因 B 完全匹配的寡核苷酸，而另外两行则含有在不同位置上的单碱基错配。通过比较正常 DNA 与肿瘤 DNA 的杂交结果来判定 LOH。如果来自正常组织的 DNA 是杂合性的（同时带有等位基因 A 和 B），则记为“带有信息的”，并且如果肿瘤 DNA 与正常相比较是纯合性的则记录为“LOH”，或者肿瘤 DNA 也是杂合性的则记录为“保留（retention）”（图 6-12）。如果对于一个 SNP，正常 DNA 是纯合性的，那么该 SNP 记录为“不带有信息的（uninformative）”，并且从分析中排除。多重 PCR 扩增应用于基于微阵列的 LOH 分析时所用的 DNA 量要比通常的基于 SSLP 的 LOH 分析时所用的 DNA 量少得多。这种使用少量 DNA 进行 LOH 分析的能力在所研究的肿瘤材料的数量很有限时是非常重要的。

因为 HuSNP 阵列基因分型的准确性是 95％～99％，建议根据几个相邻的 SNP 来寻找 LOH 区域，而不是仅仅依靠从单个 SNP 得到的信息来寻找。为此有一种简单的算法可以来确定最有可能的 LOH 区域，并且为每条染色体计算出两个最为可能的 LOH 窗口。一旦呈现 LOH 的两个连续 SNP 得到确定，就可以通过计算呈现 LOH 的 SNP 的总数，并减去呈现保留的散在 SNP 的数量，计算出最大的窗口尺寸（window size）（受到影响的染色体区域）。有一点很重要，要注意由于基因分型错误，在基因组中 SNP 位置的不恰当分配，或者是由于正常 DNA 对肿瘤 DNA 的污染，都可能出现一

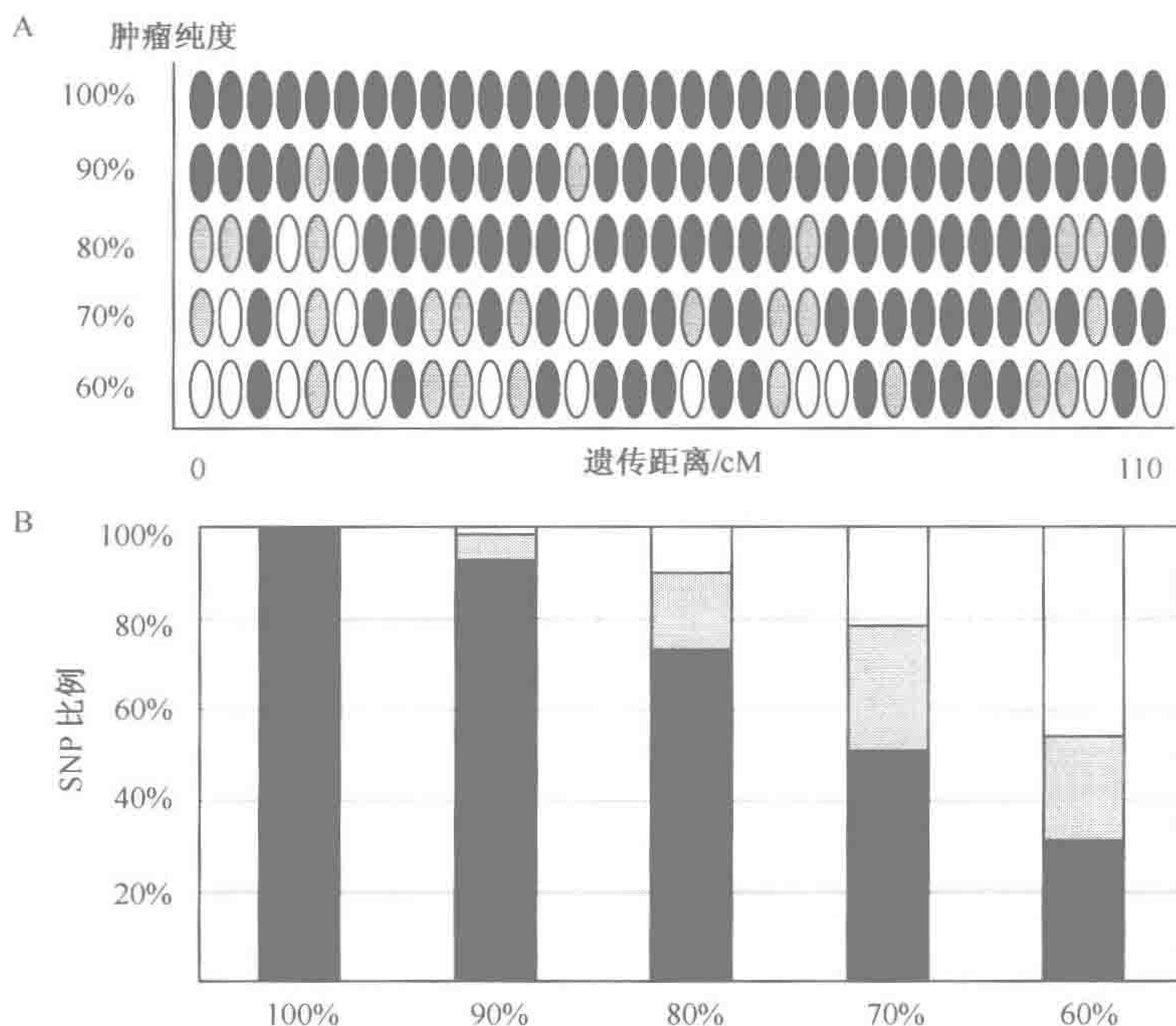


441

图 6-12 正常和肿瘤 DNA 与 SNP 阵列杂交荧光强度的典型图像

根据每个 SNP 预先确定的值，用等位基因的强度分数 A/（A+B）来决定基因型（见“实验方案”中介绍的算法）。对于每一对样品，分别给出 LOH、保留或者是不带有信息的读出（call）。WIAF-3518 是 9 号染色体上的一个 SNP 标记，WIAF-3387 与 WIAF-1765 是 3 号染色体上的 SNP

些错误的散在保留位点。当全部的 SNP 都在人类基因组上得到直接定位后，正确的 SNP 位置分配问题应该可以得到解决。因为所有的实体肿瘤（solid tumor）都有一定程度的正常组织的污染，考虑肿瘤细胞中发生 LOH 位点的检测是否被来自正常细胞的



442

图 6-13 肿瘤纯度对 SNP 判定的影响

A. 肿瘤混合实验。在相应于正常 DNA 的混合物（60%～90% 的肿瘤含量）中，对肿瘤细胞纯系 DNA（100%）中位于染色体 3p 上显示 LOH 的 SNP 进行分析。随着肿瘤 DNA 含量的下降，SNP 显示出 LOH 的变化，从 LOH 到不确定直至错误的保留。（黑色）LOH；（灰色）不确定的 LOH；（白色）杂合性保留。B. 在肿瘤混合试验中，考虑全部 130 个带有信息的样品，研究肿瘤 DNA 含量对“真实的”LOH 读出的准确性的影响。（黑色）LOH；（灰色）不确定的 LOH；（白色）杂合性保留。注意肿瘤含量为 80% 或 90% 的样品给出的结果与从 100% 肿瘤样品中得到的结果非常相似，但是在低于 80% 时，结果的质量急剧下降

DNA 掩盖是非常重要的。在实际操作中，使用 HuSNP 阵列进行基因分型时，通过区分来自两个可能的等位基因的相对信号是否落入预先确定的该特定 SNP 全部可能的基因分型的边界值之中来决定基因型。因为这些边界是根据经验为每个 SNP 设定的，一些 SNP 具有非常窄的边界而几乎不能容许污染，但是其他的却对污染非常不敏感。如果有较高水平的污染，相对等位基因信号值会位于确定的边界之外并且成为不明确的读出。如果污染水平更高，在给定的位点处肿瘤 DNA 会被错误地确定为杂合性的（SNP 确定为保留的，见图 6-13）。在实际应用中，肿瘤纯度大于 80%时就可以用 HuSNP 阵列正确地指派 LOH 区域，而低纯度会导致难以解释的结果。用 200 ng 纯度大于 80% 的肿瘤 DNA、100 ng 相匹配的正常组织 DNA 和三块 HuSNP 阵列，就可能在 2～3 d 的时间里得到单个肿瘤样品中全基因组范围内 LOH 分布的合理评价。

为了提高检测微小缺失的几率，建议选取一个包含数千个分布于全基因组的高度杂合性的 SNP 的理想集合。一种实现基于高密度微阵列进行筛查的方法是使用 SBE 标签微阵列（见方案 6 和方案 7）。HuSNP 芯片可以和更为灵活的方法，如 SBE 标签微阵列联合使用。在这种情况之下，HuSNP 筛查可以提供初步的观测，并且任何被鉴定出来的重要区域可以用具有更高分辨率的技术进行进一步的分析。

以下的方案直接改编自 Lindblad-Toh 等（2000a）；实验步骤基本上遵循 Affymetrix 为 GeneChip HuSNP 芯片所提供的方案。下面列出了主要的阶段，SNP 阵列的设计方案在图 6-14 中给出。

- 待分析 SNP 的多重扩增
- 样品的合并
- HuSNP Chip 杂交
- 数据分析

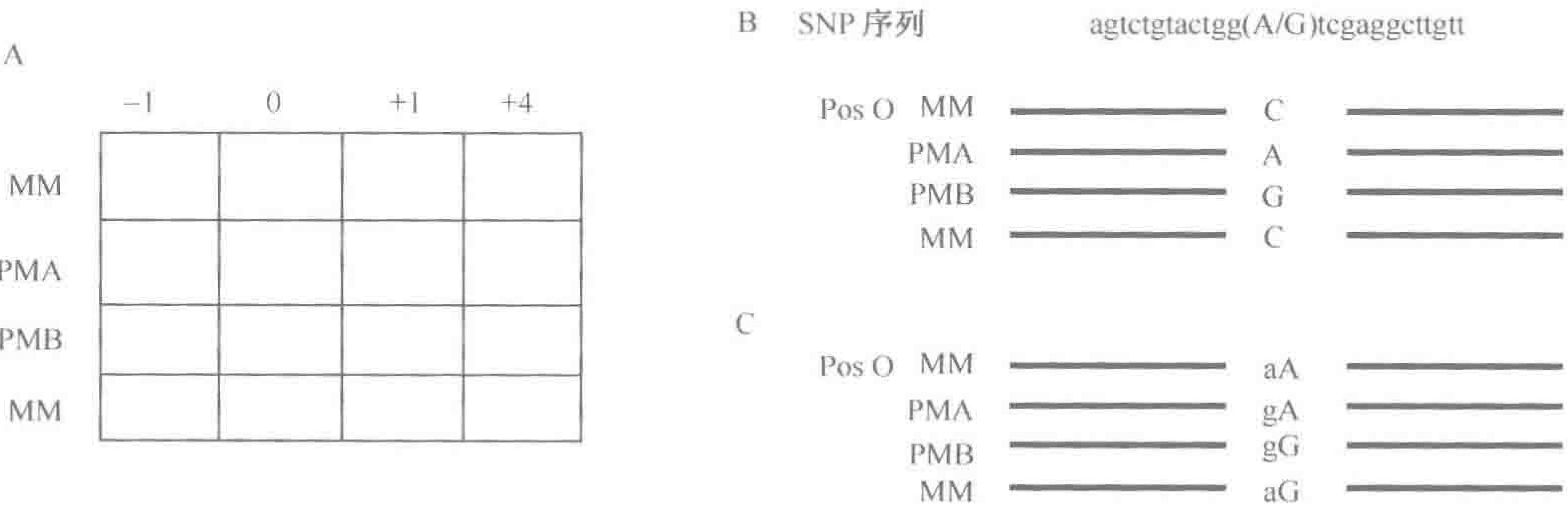


图 6-14 SNP 阵列的设计

A. 在四个位置上对每个 SNP 进行查询，使得 SNP 位于长度为 25 个碱基的寡核苷酸的中间位置（0）或者偏移 −1、+1、+4 位置。每列均有四行：错配（MM）、完全匹配 A（PMA）、完全匹配 B（PMB）、或者错配（MM）。对应于等位基因 A、等位基因 B，或者是两者（杂合子）的行会显示信号。B. 在 0 列中，SNP 碱基可能为 A 或 B，或者在两个 MM 行中是错配的碱基。C. 在其他的位置上（−1、+1 和+4），SNP 等位基因在上两行中为等位基因 A，而在下两行中为等位基因 B，并且在查询位置上两个等位基因同为匹配或错配的

材料

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

缓冲液、溶液和试剂

有关储存液、缓冲液和试剂的组成参见附录 2。
使用前将储存液稀释到合适的浓度。

10×缓冲液Ⅱ（随 AmpliTaq Gold 提供，Perkin Elmer）
ddH_2O
50×Denhart's 溶液（Sigma）
EDTA（0.5 mol/L）（GIBCO）
$MgCl_2$溶液（25 mmol/L）
链亲和素-藻红蛋白（SAPE）（Molecular Probes）
20×SSPE（BioWhittaker）
20×TBE（BioWhittaker）
TE
  10 mmol/L Tris-HCl（pH 8.0）
  0.1 mmol/L EDTA
TMACl（Sigma）
Tris-HCl（1 mol/L，pH 7.8）
Triton X-100（Mallinkrodt）
Tween-20（Pierce Chemical）
清洗溶液 A（6×SSPE-T）

30×SSPE	300 ml
100% Triton X-100	50 μl
ddH_2O	700 ml

转动试剂瓶使溶液充分混合，室温下保存。确保清洗溶液 A 在进行扫描至少一天之前配制。

清洗溶液 B（4×SSPE-T）

30×SSPE	200 ml
100%Triton X-100	50 μl
ddH_2O	800 ml

转动试剂瓶使溶液充分混合，室温下保存。确保清洗溶液 B 在进行扫描至少一天之前配制。

酶和缓冲液

AmpliTaq Gold（随 10×缓冲液Ⅱ同时提供；Perkin Elmer）

抗体

抗体(山羊)、抗链亲和素、生物素化的（0.5 mg）（Vector Laboratories）

使用前向瓶中加入 1000 μl ddH_2O 并混合。将溶液保存在 4℃下。

凝胶

琼脂糖凝胶（4%）（Nusieve 3∶1 Plus，24 孔板形式；FMC Bioproducts）

微型琼脂糖凝胶电泳装置和合适的缓冲液，溴化乙啶〈!〉，以及上样染料，1%琼脂糖。

核酸和寡核苷酸

444

DNA 序列梯（100 bp）（GIBCO）

dNTP 溶液，含有全部四种 dNTP，每种浓度均为 10 mmol/L（pH 8.0）（Pharmacia Biotech）

来自正常组织的基因组 DNA（对照 DNA）

来自肿瘤组织的基因组 DNA，用如 QiaQuick 柱进行分离

鲱鱼精 DNA（Promega）

标记引物

标记引物由用于给多重引物加尾的携带了 T3 和 T7 序列的生物素标记引物对组成。标记引物用于第二轮扩增反应，并且是随同 HuSNP 试剂盒一起提供的。

多重引物集合 1-24

这些引物集合由 T3 和 T7 加尾的引物对组成，24 个集合中的每个代表了 50～100 个遗传位点的散在集合。多重引物用于第一轮扩增反应，并且是随同 GeneChip HuSNP 试剂盒一起提供的。见 PART Ⅳ 中的表 6-7。

寡核苷酸 B1 对照（来自 GeneChip HuSNP Mapping 的试剂盒）

专用设备

管盖

离心机（吊桶式和台式）

密封膜（Microseal “A”）（MJ Research）

GeneChip Fluidcs Station 400（Affymetrix）

GeneChip HuSNP 探针阵列（Affymetrix）

GeneChip HuSNP 试剂盒（Affymetrix）

GeneChip 杂交炉 320（用于阵列）（Affymetrix）

预置为 95℃的加热槽

微量浓缩器（Microcon YM-10）（Millipore）

用于 PCR 的微孔板（96 孔，聚丙烯，V 形底）（MJ Research）

微型试管用 tough spot 胶粘标签（US Scientific）

移液器（1～10 μl 和 5～50 μl 多道）

移液器（用于转移；小内径）

移液器吸头（无菌滤芯，不含 RNase）

移液器吸头（无菌滤芯和不含滤芯的）

加长的吸头必须是尖头的，而非圆头的，以便于向探针阵列上加液（US Scientific）。

多孔板（1 ml 深孔滴定板）（Beckman）

多孔板（1 ml 深孔，圆底的）

扫描仪（HP Gene Array 或 Affymetrix 的 Agilent Scanner）

扫描软件（Affymetrix GeneChip 3.1）

热循环仪

热循环仪要按照需要的扩增方案进行编程（MJ Research DNA Engine 或 Tetrad；或者 Perkin Elmer 2400、9600 或 9700）。

带管盖的试管（0.2 ml 薄壁的，8 连排）（MJ Research）

试管（1.5 ml 微量离心管，不含 RNase）（US Scientific）

真空过滤装置（0.20 μm 或 0.45 μm，容量为 1L）（Corning）

445

方法

待检测 SNP 的多重扩增

对于三份要进行 LOH 检测的 DNA，用 HuSNP 芯片对每份样品进行基因分型：正常的对照组织 DNA 和包含重复的肿瘤 DNA。对于这些样品中的每一份，需要检测 1494 个 SNP 位点，并且分为 24 个多重 PCR 引物的集合。

1. 进行第一轮多重 PCR 扩增。

   a. 每份基因组 DNA 用 TE 稀释到 4 ng/μl，并分别将正常和肿瘤 DNA 标记为 $N_{individual}$、$T_{individual}$a 和 $T_{individual}$b。注意肿瘤 a 和 b 是相同样品的重复。为每份样品标记一支微量离心管，并向每支试管中加入 35 μl 相应的基因组 DNA。

   b. 按照下述在微量滴定板中建立多重扩增反应：

      i. 按照以下方案为每种样品标记一块微孔板：

	1	2	3	4	5	6	7	8	9	10	11	12
A	1	9	17	1	9	17	1	9	17			
B	2	10	18	2	10	18	2	10	18			
C	3	11	19	3	11	19	3	11	19			
D	4	12	20	4	12	20	4	12	20			
E	5	13	21	5	13	21	5	13	21			
F	6	14	22	6	14	22	6	14	22			
G	7	15	23	7	15	23	7	15	23			
H	8	16	24	8	16	24	8	16	24			
样品	正常 1			肿瘤 1a			肿瘤 1b					

ii. 向每四分之一板中加入 3 μl 每种多重引物集合 1～24。

iii. 为每个 DNA 样品准备足够 28 个反应的预混试剂（24 个反应孔以及额外的混合液以防少量移液误差）。

试　剂	1 反应	1 样品（28×）
10×缓冲液 Ⅱ	1.25	35.0
25 mmol/L $MgCl_2$	2.5	70.0
10 mmol/L dNTP 溶液	0.625	17.5
5 单位/μl AmpliTaq Gold	0.25	7.0
H_2O	3.625	101.5
总体积/μl	8.25	231.0

iv. 向每支微量离心管（步骤 1a）中加入 231 μl 预混试剂并混合。

v. 对于每四分之一板中的 24 个加样孔，向每孔中加入 9.5 μl 含有引物集合的预混试剂。用盖子封好微孔板。

c. 将微孔板置于热循环仪中，并按照以下的变性、复性及聚合时间和温度进行扩增反应：

446

循环次数	变　性	复　性	延伸（聚合）
1	95℃ 10 min		
30	95℃ 30 s	55℃ 1 min，自动 +0.2℃/循环	72℃ 30 s
5	95℃ 30 s	58℃ 1 min	72℃ 30 s
1			72℃ 7 min

这些时间是为本方案中介绍的特定设备和反应而优化的，需要对条件进行修改以适合于其他类型仪器和反应。

2. 配制多重扩增产物的 1∶1000 稀释液：

a. 向 96 孔板中加入 1 ml ddH_2O。准备多孔板，多孔板数目和第一轮 PCR 中所使用的数量相同，加样孔的布局与步骤 1 中所述相似。用 Microseal 塑料封口膜密封多孔板，直到进行步骤 b。

b. 从步骤 1 的扩增反应产物中各取出 1 μl 多重扩增产物加入到准备好的多孔板的相应加样孔中，得到多重扩增产物的 1∶1000 的稀释液（继续直到所有产物均被稀释）。用 Microseal 塑料封口膜在多孔板上方密封，并用滚筒将塑料膜压平，使其密封严实。轻轻涡旋振荡样品。

c. 在室温下，将多孔板在大的吊桶式离心机内以 1000 *g* 离心 2 min，以回收封口膜上的液体。

3. 进行第二轮多重 PCR。

a. 按照下述步骤在微量滴定板中建立第二轮多重扩增反应体系：

i. 使用与步骤 1 中介绍的第一轮扩增反应中使用的多孔板相同的样式，标记第二轮的多孔板。

ii. 从装有稀释的多重扩增产物的深孔板上除去封口膜，移取 2.5 μl 稀释的样品到新准备的多孔板上对应的加样孔内。

iii. 按照下列顺序加入以下试剂，为每种 DNA 样品准备足够 28 次反应的预混标记试剂：

试　剂	反应（1×）	1 样品（28×）
10×缓冲液 Ⅱ	2.5	70
25 mmol/L $MgCl_2$	4.0	112
10 mmol/L dNTP 溶液	1.0	28
10 μmol/L 生物素标记的 T7 引物	2.0	56
10 μmol/L 生物素标记的 T3 引物	2.0	56
5 单位/μl AmpliTaq Gold	0.5	14
ddH_2O	10.5	294
总体积/μl	22.5	630

iv. 向微孔板的每孔中加入 22.5 μl 预混标记试剂。用盖子密封微孔板。

b. 将微孔板置于热循环仪中，并按照以下的变性、复性及聚合时间和温度进行扩增反应：

循环次数	变　性	复　性	延　伸
1	95℃ 10 min		
40	95℃ 30 s	55℃ 90 s	72℃ 30 s
1			72℃ 7 min

这些时间是为本方案中介绍的特定设备和反应而优化的，需要对条件进行修改以适合于其他类型仪器和反应。

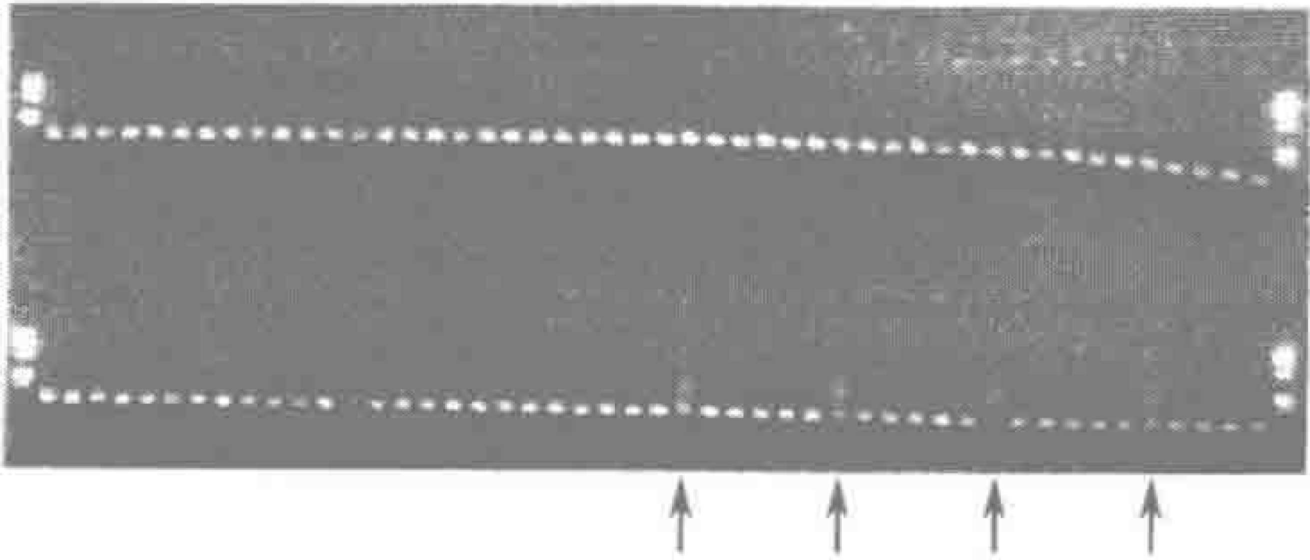


图 6-15　肿瘤 DNA 纯度对 SNP 评分的影响

每个集合显示为一条特异性的条带。箭头标记的集合显示模糊，并且可能要重新扩增，特别是正确大小的产物发生缺失时（左起第三个箭头）

4.（可选的）在 1%琼脂糖凝胶上检查每个集合的扩增产物。

a. 将 2 μl 每孔产物与琼脂糖上样染料混合，并用 1%的琼脂糖凝胶电泳分离产物。

b. 对凝胶进行分析。确保每个集合都具有大致相同大小（50～100 bp）与染色强度的散布条带。

c. 对于任何没有产物的样品，对该特定的样品/PCR 集合组合重新进行步骤 1～4。需要根据单个反应的次数，而不是样品数来计算所需的混合溶液的量。

通常，PCR 集合扩增良好，但是如果后来的故障检查阵列的结果变得必不可少时，凝胶电泳可以提供有用的信息（图 6-15）。

样品合并

重要事项：必须合并浓缩每种 DNA 样品的产物。避免污染以及样品间的混合。

5. 将产物短时离心，从盖上除去液体。建议合并 PCR 产物时，只打开同一给定 DNA 样品的 24 个扩增集合的盖子。将 24 个加样孔中的产物取出并合并于一个洁净的、标记好的微量离心管内。
6. 浓缩样品
   a. 为每份样品标记一个 Microcon YM-10 过滤装置。确保过滤柱与第二支收集管都进行了标记。
   b. 将合并的样品加入置于第一支收集管内的绿色过滤装置中。
   c. 室温下，在固定角度转子的台式离心机中，以 13 000 *g* 离心 26 min 收集样品。
   d. 小心地倒转，将绿色过滤装置放入第二支带有标记的收集管内，并弃去第一支收集管。到这里体积大约为 60 μl。最好不使用移液器测量体积，以免吸头损坏滤膜。向一个空的过滤装置中加入 60 μl 的样品来进行比较。
   e. 室温下，在台式离心机中以 3000 *g* 离心 3 min 来收集最终的样品。
   f. 测量收集到的产物的体积。如果在 60～90 μl，继续进行步骤 g。如果体积少于 60 μl，加入 ddH_2O 至 60 μl。如果体积超过 90 μl，重复步骤 c～f，并且只需要离心 2 min。非常重要的是不要将样品离心至滤膜上干燥。
   g. 当所有样品达到正确的体积后，将它们在－20℃下保存直至继续杂交步骤。

448

HuSNP 芯片杂交

7. 按照下述步骤进行杂交反应：
   a. 为每份样品标记一支微量离心管。向每支离心管中加入 30 μl 浓缩的合并产物（来自步骤 6）。
   b. 在离心管内依次加入以下试剂配制杂交混合物（该配方足够一份样品用）：

杂交混合物	1×/μl
5 mol/L TMACl	81
对照寡核苷酸 B1	1.35
1 mol/L Tris-HCl (pH 8.0)	1.35
1% Tween-20	1.35
0.5 mol/L EDTA (pH 8.0)	1.35
10 mg/ml 鲱鱼精 DNA	1.35
50× Denhardt's 溶液	13.5
ddH_2O	3.75
总体积	105

注意：对照寡核苷酸对于扫描仪进行识别阵列的类型和对齐阵列的边角是必需的。

c. 向每支装有合并样品的微量离心管（步骤 a）内加入 105 μl 杂交混合物，并在微型离心机内短暂离心以混合液体。

d. 为了使样品变性，将微量离心管在 95℃的热槽内温浴 10 min。

e. 在样品变性的同时，使用 Tough-spot 黏性标签对 GeneChip HuSNP 探针阵列进行标记。保证所有芯片来自同一批次，并记下产品批号。将一个 P-10 吸头插入每块芯片上的两个通气孔之一。

f. 将杂交反应在阵列的杂交腔内进行就不会产生过多的气泡，这一点非常重要。用细长的 300 μl 吸头将每种变性的样品加入到芯片上：缓慢地将吸头插入阵列下方的通气孔中。然后在保持芯片垂直的情况下缓慢地将液体推入芯片中，使得空气向上排出，通过芯片窗口一侧观察液体的注入。

g. 从上方的通气孔中取出 P-10 吸头。在过夜的杂交步骤中，将阵列腔密封以避免蒸发。如果有蒸发发生，杂交混合物中的盐分会沉淀出来，从而使阵列变得无效。用 Tough-spot 黏性标签密封阵列上的两个通气孔。

h. 将芯片放在杂交炉内的旋转支架上，使其平衡，以 45 r/min 的速度旋转过夜。

i. 在 45℃下将阵列杂交过夜（或进行 16 h）。

449

8. 对杂交后的阵列清洗和染色：

a. 使用清洗溶液 A 和 B 灌注 Affymetrix Fluidics Station。

b. 在 50 ml 的 Falcon 管中混合下面的染色溶液，并短暂离心以混合。

染色溶液	1×/μl
20×SSPE	150
50×Denhardt 溶液	10
1% Tween-20	5
1 mg/ml SAPE	25
ddH_2O	305
总体积	495

c. 将染色溶液按照 495 μl 分装到带有标记的微量离心管中。将试管置于冰上，用铝箔覆盖，因为链亲和素-藻红蛋白（SAPE）对光敏感。

d. 从杂交炉中取出恰好能够装满洗涤工作站数量的芯片（通常为四块芯片/洗涤工作站）。撕去覆盖芯片上通气孔的 Tough-spot 黏性标签。

e. 每份样品使用一支装有染色混合液的微量离心管（步骤 8c）。向每支微量离心管中加入 5 μl 的 0.5 mg/ml 生物素化抗链亲和素抗体，加液时向试管内壁上滴加一滴液体。快速涡旋振荡混匀以避免 SAPE 与抗体发生沉淀。

在染色溶液中同时加入 SAPE 和生物素化抗链亲和素抗体可以通过信号放大来增强信号微弱位点处的信号强度，因为有多个 SAPE 和抗体分子结合到每个生物素化的 PCR 产物中。

f. 在 Affymetrix Fluidics Station 中立即进行清洗/染色步骤。根据仪器指示，将芯片与染色试剂瓶先后放入洗涤工作站。用下面的条件设置清洗与染色方案：

溶 液	温 度	清洗次数	每次清洗的混合次数
清洗溶液 A	25℃	2	2
清洗溶液 B	35℃	6	5
染色溶液	25℃	1800 s	
清洗溶液 A	25℃	6	4

g. 清洗结束后，将芯片在黑暗中保存并尽快扫描。

室温下，芯片可以稳定保存 12 h，但是建议在清洗结束后立即进行扫描。

h. 按照需要的次数重复步骤 d～g 以处理剩余的芯片。

9. 使用 Affymetrix GeneChip 3.1 软件（或可获得的更新版本）在 Affymetrix Agilent 扫描仪中扫描 HuSNP 芯片。为了抵消来自强信号位点的信号放大，同时用 530 nm 和 570 nm 对阵列进行扫描——尽管 570 nm 是最佳的发射波长，在 530 nm处的信号强度仅约为 5%。因此，强信号位点可以在 530 nm 处读取，并将其信号外推至 570 nm 处。
   a. 制备标记有使用日期和样品名称的实验文件，例如，L010601N、L010601Ta 和 L010601Tb，以确保探针阵列的类型设置为 HuSNP。
   b. 用 530 nm 和 570 nm 波长扫描芯片（检查扫描仪的选项）。
   c. 将芯片于室温下黑暗中保存，直到最初的分析完成。

   步骤 10 a～c 与扫描最好在同一天进行。

450

数据分析

数据分析可以分成两个部分：① 每份样品的基因分型；②单个位点上 LOH 的分析和筛查的每个个体染色体区域上的 LOH 分析。基因分型使用 Affymetrix 软件自动进行，而 LOH 的分析需使用定制的软件脚本。

10. 对每份样品进行基因分型：
    a. 将网格与芯片上对照寡核苷酸的正方形轮廓对齐。通过生成 .cel 文件从每个叠加的网格中提取信息。
    b. 运行分析模块，产生含有每个标记的基因型的一个芯片文件。
    c. 运行报告工具，检查读出率要大于 75%，每块阵列上各个波长的相关系数要大于 98%，以此来检查扫描的质量。

    任何低于这些数字的阵列都不能用于分析。关于阵列失败原因的讨论在方案最后疑难问题解答的信息栏中给出。

11. 进行基本的 LOH 分析：对于每份样品，通过对比正常对照和肿瘤重复的读出结果，判定每个 SNP 为杂合性缺失（L）、杂合性保留（R）或者无有效信息/不确定（U）。通过指定恰当的状态对每份样品进行判定：
    - LOH＝正常中的任何 SNP 都显示为两个等位基因（*AB*），并且肿瘤重复中至少有一个只有一个等位基因（*AA* 或 *BB*）。
    - 保留＝正常中的任何 SNP 都显示为两个等位基因（*AB*）；并且或者是肿瘤重复中两个均为两个等位基因（*AB*），或者有一个且另一个肿瘤重复为“无信号”。

- 不确定＝正常中的任何 SNP 显示为仅有一个等位基因（*AA* 或 *BB*）或者无信号；或者正常中有两个等位基因（*AB*）且在两个肿瘤重复中均为“无信号”。

12. 累加 LOH 和保留的评分得出具有有效信息的 SNP 的数量。用 LOH 的数量除以带有效信息 SNP 的数量，计算出每个样品的 LOH 比例。
13. 指派 LOH 区域：使用该方法可以抵消偶尔出现的基因分型和作图位置错误，并且能对关键区域提供一个粗略的估计（图 6-16）。
    a. 从表 6-8 中获取每个 SNP 在图谱中的位置（PART Ⅳ，SNP 序列：作图数据），并创建所分析的 SNP 的图谱。该文件中全部 1081 个 SNP 均使用

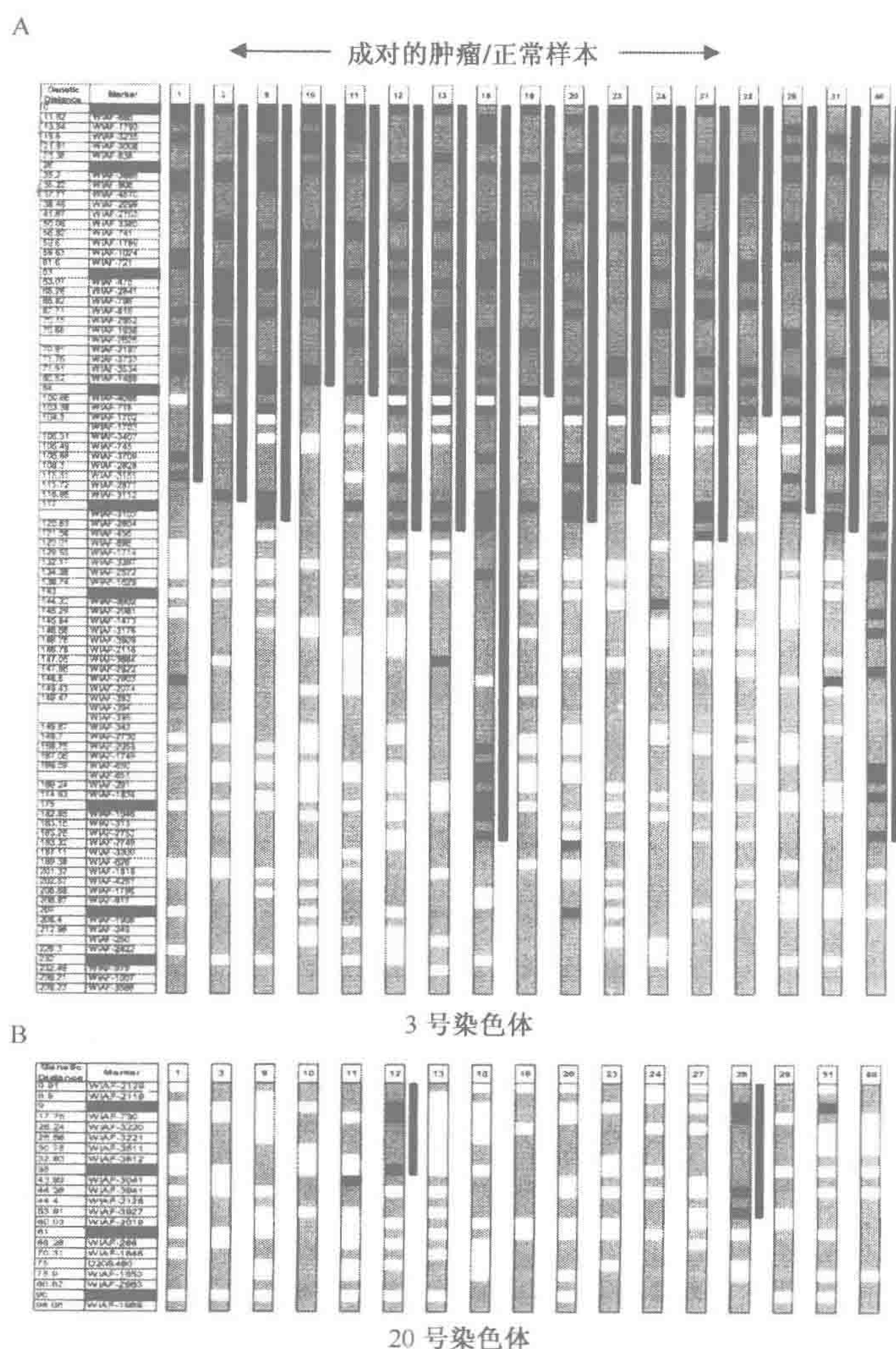


图 6-16　3 号（A）和 20 号（B）染色体作图示例，带有预测的 LOH 窗口

“遗传距离”对于 SSLP 是真实的遗传距离，而 SNP 是推断的遗传距离，这是基于辐射杂种系的作图位置。黑色的标记为 SSLP。每对标记/样品注释如下：（黑色）LOH；（白色）保留；（灰色）无有效信息。根据本方案所述计算出的预测的 LOH 窗口，在每份样品的右侧用黑色方块标记

TNG 辐射杂种版 http://www-shgc.stanford.edu/RH/index.html 进行作图，并且和 Genebridge4 辐射（Walter et al. 1994）杂种版的结果一致，SNP 在这两个杂种版中都已经得到了作图。

b. 忽略所有不带有信息的位点，并按照下列步骤来确定 LOH 的近似区域： 452

 i. 从短臂的顶端开始筛查每条染色体，鉴别出最先发生连续缺失的两个标记所在位置。

 ii. 初始化一个工作计数器，σ 的初始值设置为 2（表明两个连续的标记呈现缺失）。

 iii. 每个后续的缺失位点，则计数器为 +1，而每个后续的保留位点，计数器为 −1。

 iv. 将 σ 值最大的两个区域定义为 LOH 区域。

如果靠近染色体任何一端的具有有效信息的 SNP 被认为是窗口的一部分，那么余下的染色体末端不带有效信息的 SNP 也被认为是窗口的一部分。图 6-16 中显示了从这个分析输出的最终数据类型的一个例子。

疑难问题解答

阵列的读出率低或者相关因子差

正如上面所注意到的，检测每个阵列的相关系数和读出率是非常重要的。因为根据相关系数将 530 nm 的读出数据外推至了 570 nm，所以，如果相关系数小于 0.98，就会冒基因分型读出不正确的风险。低的读出率，或者读出率和相关系数两者均低的情况偶尔会发生，在这种情况下，阵列图像显示出比平均水平或多或少要亮的单元。读出率或相关系数差的问题常常与阵列的批次有关，并且可能与阵列的合成质量相关联。

- 如果检测到这些问题，就通过目测阵列图像来将具有正常相关系数和读出率的阵列和存在缺陷值的那些阵列进行比较。在比较阵列时，确保图像显示中采用的信号放大程度相同。
- 如果杂交信号低而且读出率低，用一块新的阵列再次杂交样品。如果样品仍然失败，重新进行扩增。
- 如果读出率过低，相关系数低，但是信号强度高（阵列非常明亮），尝试只用 5～15 μl 样品进行杂交（重做步骤 7～9）。在步骤 5a 中加水至 30 μl 来调整样品的体积。
- 如果许多样品的结果差，并且扩增步骤看上去工作良好（可选择的步骤 4：产物的凝胶电泳），该批次阵列可能存在缺陷，在新批次的阵列上重新进行实验。

PART Ⅲ 基于微阵列的 DNA-蛋白质相互作用检测：微阵列上的染色质免疫沉淀 453

Vishwanath Iyer（得克萨斯大学，奥斯汀，得克萨斯州，78712-1095）

导言

基因组相关联的细胞过程，如转录、DNA 复制、DNA 修复、重组和染色体分离，

都涉及基因组 DNA 与蛋白质的相互作用。鉴定每种 DNA 结合蛋白的结合部位或位点对于在分子水平上研究这些过程是根本性的。很多方法能够鉴别一种特定蛋白质的靶 DNA 序列，但是最近染色质免疫沉淀（chromatin immuneoprecipitation）方法得到了广泛的应用。该方法依赖于特异性抗体使感兴趣的蛋白质与和它结合的 DNA 序列被同时免疫沉淀。应用甲醛进行共价交联能够检测到体内特定的相互作用，并且已经用于酵母（Strahl-Bolsinger et al. 1997）、果蝇（*Drosophila*）（Orlando and Paro 1993），以及哺乳动物细胞内（Boyd and Farnham 1999）的 DNA-蛋白质相互作用作图。

较近一段时期以来，该方法和微阵列杂交相结合，能够在基因组的范围内对酵母中 DNA-蛋白质的相互作用进行作图（Reid et al. 2000；Ren et al. 2000；Iyre et al. 2001）。可以制备基本上能够代表每个酵母启动子的微阵列（见 PART Ⅰ，方案 6）使得该方法成为一种特别有力的方法。与大多数牵涉到点样 DNA 微阵列的实验一样，该方法应用了双色荧光杂交的策略。使用一种荧光染料标记特异性免疫沉淀的 DNA（方案 9）。第二种染料用于一个标记参考样品，可以简单地由基因组 DNA 构成，同时进行扩增和标记，或者是从平行的免疫沉淀反应中得到的，用作对照的 DNA。用带有一个相关基因缺失的酵母菌株和缺少融合的表位标签的抗体，以及不相关的抗体，或者根本没有抗体，来进行该免疫沉淀。在诸如这些对照的情形下，进行免疫沉淀反应后回收的 DNA 最有可能在大体上代表了基因组 DNA，并且相互之间以及和参考样品之间是可以替换的。但是，宁愿使用各种对照或者参考样品的组合进行微阵列杂交，也不在表面上接受这一假设。

差异标记的参考样品和免疫沉淀样品同时与代表全部酵母基因间隔区序列的微阵列杂交之后（方案 10），每个阵列单元的荧光比值正比于免疫沉淀样品中该位点的富集，因此也就指示了蛋白质与该位点的结合程度。

454 天然蛋白质或者融合到蛋白质的表位标签的单克隆和多克隆抗体已经被成功地应用于该方法之中。表位标签的优势在于相同的免疫沉淀条件可以应用于许多的蛋白质，而且对于大多数常用的表位标签，如流感病毒血凝素（influenza virus hemagglutinin, HA）或者 Myc，可以获得那些已被证明适用于免疫沉淀反应的高亲和力抗体。但是，要使得该方法能够成功，融合表位标签形式的蛋白质在体内必须是有功能的，并且在交联之后，表位必须以 DNA 结合的形式暴露出来。

455 方案 10

应用阵列对酵母基因组上 DNA-蛋白质相互作用作图

本方案介绍了基于微阵列的酵母中 DNA 结合蛋白分布作图的步骤。整个过程（不包括基因芯片的制备）可以在 3～4 天内完成。通过在培养物中接种适量的细胞，酵母细胞可以过夜培养，并且在第 1 天的上午进行交联和收获。如果与抗体孵育的时间很短（4 h），那么解交联反应就能够过夜进行。第 2 天的工作是纯化免疫沉淀的 DNA 和用氨基-烯丙基修饰的核苷酸进行标记反应。与 Cy 染料的耦联、微阵列杂交（6 h）和扫描可以在第 3 天完成。每次仅用一部分免疫沉淀的 DNA 进行微阵列杂交。三分之一的材料用

于氨基-烯丙基标记，并且得到的DNA只有一半用于耦联反应和微阵列杂交。因此每次免疫沉淀反应（immunoprecipitation，IP）提供了足够至少6次微阵列杂交的材料。

材料

注意：标记有〈!〉的材料的正确处理方法参见附录3。

缓冲液、溶液和试剂

有关储存液、缓冲液和试剂的组成参见附录2。

使用前将储存液稀释到合适的浓度。

苯甲脒

牛血清白蛋白（BSA）（0.5 mg/ml 储存液）

氯仿〈!〉

二硫苏糖醇（DTT）（0.1 mol/L）〈!〉

EDTA

乙醇

甲醛（37%）〈!〉

甘油

甘氨酸（2.5 mol/L）

糖原（20 mg/ml）

N-2-羟乙基哌嗪-*N*′-2-乙磺酸-氢氧化钾（HEPES-KOH）（1 mol/L，pH 7.5）

IP洗脱缓冲液：

- 50 mmol/L Tris-HCl（pH 8.0）
- 10 mmol/L EDTA
- 1% SDS〈!〉

IP清洗缓冲液Ⅰ

- 50 mmol/L HEPES-KOH（pH 7.5）〈!〉
- 150 mmol/L NaCl
- 1 mmol/L EDTA
- 0.1% 脱氧胆酸钠
- 1% Triton X-100

IP清洗缓冲液Ⅱ

- 50 mmol/L HEPES-KOH（pH 7.5）〈!〉
- 500 mmol/L NaCl
- 1 mmol/L EDTA
- 0.1% 脱氧胆酸钠
- 1% Triton X-100

IP清洗缓冲液Ⅲ

- 10 mmol/L Tris-HCl（pH 8.0）

456

250 mmol/L LiCl
1 mmol/L EDTA
0.5% 脱氧胆酸钠
0.5% Nonidet P-40

KCl

LiCl

裂解缓冲液

50 mmol/L HEPES-KOH (pH 7.5)
150 mmol/L KCl
1 mmol/L EDTA
10% 甘油
0.1% Nonidet P-40

$MgCl_2$ (25 mmol/L)

N-羟基琥珀酰亚胺酯耦联的 Cy3 和 Cy5

Nonidet P-40 (10% *m/V*)

酚：氯仿 (1：1) 平衡的 〈!〉

苯甲基磺酰氟 (PMSF)〈!〉(100 mmol/L 储存液，溶于异丙醇)

蛋白酶抑制剂 (在使用前加入)

1 mmol/L 苯甲脒
1 mmol/L 偏亚硫酸氢钠
0.2 mmol/L PMSF
1μg/ml 抑胃酶肽

蛋白质 G-Sepharose

SDS (10%)〈!〉

醋酸钠 (3 mol/L，pH 5.2)

脱氧胆酸钠 (10% *m/V*) 溶于 10 mmol/L 的 HEPES-KOH (pH 8.0)

偏亚硫酸氢钠

TE (pH 8.0)

Triton X-100

酶和酶缓冲液

胃蛋白酶抑制剂

蛋白酶 K

RNase A

457 抗体

感兴趣的表位标签或蛋白质的一抗

细胞

Saccharomyces cerevisiae 菌株

专用设备

玻璃珠（500 μm）

带 O 型圈的微量离心管（旋盖的）

Mini-Beadbeater-8 快速型珠磨式组织研磨器（Biospec Products Inc.）

　　可选的，见步骤 6。

Nutator（Clay Adams Inc.）

　　可选的，见步骤 10。

QIAquick PCR 纯化试剂盒（QIAGEN）

注射器（6 ml，带有 25 号针头）

针式超声波发生器

方法

交联和免疫沉淀

1. 培养200 ml 酵母细胞至 OD_{600} 约为 0.6～0.8。

   在完全培养基中培养的健康酵母菌株应该处于培养的对数期。本方案中试剂的用量是基于这一假设的。200 ml 的培养物通常可以提供足够供给 4 次免疫沉淀反应的材料。对于不同的菌株、培养条件、细胞密度和培养体积、试剂的用量可以作相应的调整。

2. 直接向培养物中加入 5.6 ml 37%的甲醛，使终浓度为 1%，以交联 DNA 和蛋白质。在室温（26℃）下继续孵育 15～30 min，振荡速度 100 r/min。

   与甲醛孵育 30 min 看来足以完全交联大多数 DNA-结合蛋白。值得去尝试一系列的交联时间，尤其是对于那些与 DNA 相互作用的机制还不十分清楚的蛋白质。但是，过度交联对于后续的抗体结合和免疫沉淀步骤可能是有害的。对给定的一系列实验，交联时间应该保持恒定。

3. 加入 11 ml 浓度为 2.5 mol/L 的甘氨酸（至终浓度为 125 mmol/L）以猝灭交联反应。继续在室温下将培养物振荡 5 min。

4. 在 4℃下以 1500 *g* 离心 5 min 收集细胞。将细胞在冰冷的 PBS 中洗涤两次。在洗涤过程中，将细胞合并到一支 15 ml 的圆锥形底试管中。

5. 加入 2 ml 冰冷的裂解缓冲液，并用 1 ml 的移液管上下吹打重悬细胞。分别向两支 2 ml 带旋盖，并加有橡胶 O 型圈的微量离心管内各加入 1 ml 的细胞悬液。向每支试管内加入约 1 ml 的玻璃珠（500 μm）并旋紧管盖。

   用一支洁净的微量离心管，外侧根据 1 ml 水的液面位置进行标志，玻璃珠可以在相同的试管中据此进行量取。

   将管盖旋到装有玻璃珠和细胞悬液的试管之前，确保没有玻璃珠黏附在管口或 O 型圈上。

6. 在 4℃下以最高速度涡旋振荡，最大可能地破碎细胞。

   因为抽提效果取决于细胞破碎的程度，充分进行本步骤非常关键。在普通的旋转振荡器中，完全破碎需要在 4℃下进行 30 min。如此长时间手持试管置于涡旋振荡器上是不切实际的。需要一个可以夹持多支试管在适当位置上的夹具。

   一个较好的选择是购买一台 Mini-Beadbeater-8 快速型珠磨式组织研磨器（Mini-Bead-

beater-8，BioSpec Products，Inc.），它能够更强烈地进行试管振荡并显著缩短破碎时间。如果使用 Beadbeater，以最大速度分 4 次振荡试管，每次 1 min，振荡期间将试管放回冰上 2 min 以防止管内物质升温。

7. 将装有 25 号针头的 6 ml 注射器针筒放入置于冰上的 15 ml 圆锥形试管中。将细胞破碎管内的物质，包括玻璃珠，倒入注射器内，使抽提液流进收集管中。加入 0.5 ml 裂解缓冲液洗涤留在注射器中的玻璃珠，并用注射器柱塞加压使剩余的液体全部流出。如果针头被堵，将柱塞稍稍提升，倒转注射器，轻弹一会儿。如果针头彻底堵塞，可以更换。
8. 用超声波处理以剪切交联抽提物中的 DNA。用超声波发生器，将探针浸入装在 15 ml 试管内的抽提液中。超声处理 3 个循环，每次 30 s，能量设定为 5，占空比 50%。每次超声处理之间将抽提物放回冰上 2 min 以保持低温。

   超声处理将交联的 DNA 随机剪切为长度 300～1000 bp 的片段。在步骤 9 之后，用少量抽提物（10 μl）进行琼脂糖凝胶电泳，可以检查超声处理的程度。

9. 将经过剪切的交联染色质转移到微量离心管中，4℃下在微型离心机中以最大速度离心 10 min，除去细胞碎片。仔细将上清液转移到新的微量离心管内。

   抽提物的体积要刚好在 2 ml 以上，这应该足够供给至少 4 次免疫沉淀反应。

   如果有需要，稀释抽提物以用于更多的免疫沉淀反应，如使用不同浓度的抗体，或者是不加抗体或使用免疫血清的模拟免疫沉淀反应，或者用于不同靶蛋白的免疫沉淀。

10. 对于每次免疫沉淀反应，向 500 μl 经过剪切的交联抽提物中加入特异性的一抗。

    有时需要根据经验来确定每种待测蛋白所使用的最佳抗体浓度。但是，通常每次免疫沉淀反应可使用 1∶100 或 1∶50 稀释，或者是 2～5 μg 的抗体。免疫沉淀反应于 4℃下缓慢摇动或者振荡，孵育 4 h 至过夜。使用 Nutator 振荡器对本步骤以及后续步骤而言是最为理想的。

11. 向抗体复合物中加入 50 μl 预先用裂解缓冲液平衡的浓度为 50%的蛋白质 G-琼脂糖珠悬液。用尖端切平的 200 μl 吸头移取蛋白质 G 小珠。将悬液在 4℃下缓慢摇动或者振荡孵育 2 h。

459

12. 在室温下将悬液以 1000 *g* 离心 15 s。小心吸出上清液，避免搅动凝胶小珠。可以在珠子中留少量的溶液。
13. 在室温下每次用 1 ml IP 清洗缓冲液 I 洗涤小珠 2 次，每次 5 min。每次洗涤后在室温下以 1000 *g* 离心悬液 15s。
14. 在室温下用 1 ml IP 清洗缓冲液 Ⅱ 洗涤小珠 1 次，时间 5 min。然后在室温下以 1000 *g* 离心悬液 15 s。
15. 在室温下用 1 ml IP 清洗缓冲液 Ⅲ 洗涤小珠 1 次，时间 5 min。然后在室温下以 1000 *g* 离心悬液 15 s。
16. 在室温下用 1 ml TE 洗涤小珠 1 次，时间 5 min。这个最后的清洗步骤之后，尽可能地吸去清洗溶液，而不要搅动小珠。

    前面的洗涤步骤对免疫沉淀反应的效果会有影响。可以通过改变清洗缓冲液中盐和去污剂的浓度，并分析它们在洗脱步骤之后是如何影响已知靶位点的特异性富集来优化清洗步骤的严格程度。如果需要，在 IP 清洗缓冲液 Ⅱ 和 Ⅲ 中使用低浓度的盐和（或）去污剂可以减低严格度。但是，重要的是整个洗涤和洗脱步骤要进行得快速而且稳定

可靠。

17. 向小珠团块内加入 100 μl IP 洗脱缓冲液来洗脱免疫沉淀的物质。将悬液混合，并在 65℃下孵育 15 min。在微型离心机中以最大速度将悬液离心 2 min，并将上清液转移至新的试管中。向小珠中加入 50 μl 洗脱缓冲液并重复洗脱步骤。合并洗脱液（约 150 μl）。
18. 在加热槽中将洗脱液于 65℃下加热至少 6 h 至过夜以进行解交联。在此过程中要确保试管的管盖已盖严，并密封好。
19. 短暂离心试管，然后加入 150 μl TE。加入 1 μl 浓度 20 mg/ml 的糖原和最终浓度为 100 μg/ml 的蛋白酶 K。将溶液在 37℃下孵育 2 h。
20. 加入 300 μl 的酚：氯仿（1：1）。将悬液涡旋振荡 10s，随后在室温下于微型离心机中以最大速度离心 10 min。将水层转移到新的试管中。向有机相中加入 100 μl TE 并涡旋振荡，与前面同样通过离心回收水相。合并水相抽提物（约 400 μl）并用等体积的氯仿抽提溶液 1 次。
21. 加入 3 μl 浓度为 0.5 μg/μl 的不含 DNase 的 RNase。将溶液在 37℃下孵育 30 min。加入醋酸钠（pH 5.2）至浓度为 0.3 mol/L 和 2 倍体积的乙醇。将含乙醇的溶液在－20℃下放置至少 1 h。
22. 4℃下，在微型离心机中以最大速度离心 30 min 来回收沉淀的 DNA。用 500 μl 70%的乙醇洗涤团块 1 次。将团块在空气中干燥，并将 DNA 团块重悬于 25 μl TE 中。将免疫沉淀的 DNA 保存在－20℃下。

方案 11

460

酵母基因组的 DNA-蛋白质相互作用作图：DNA 扩增及标记

本方法用于扩增方案 10 中制备的免疫沉淀的 DNA。扩增分成两个阶段进行：在 A 轮中，以免疫沉淀的 DNA 为模板，用一条部分简并引物非特异性地引导两个 DNA 合成循环。在 B 轮中，用特异的 PCR 引物和氨基-烯丙基标记的核苷酸作为底物，A 轮的双链在 B 轮中得到指数扩增。由于在 A 轮中使用了能够扩增多种来源 DNA 的简并引物，因此避免外源序列对免疫沉淀的 DNA 的污染非常重要。为了防止出现问题，要使用带气溶胶滤芯的吸头，在层流通风橱内戴上手套进行操作，并且要避免产生气溶胶。

材料

缓冲液和溶液

有关储存液、缓冲液和试剂的组成参见附录 2。

使用前将储存液稀释到合适的浓度。

氯化镁（25 mmol/L）

磷酸盐缓冲液（PBS）

碳酸氢钠缓冲液（1 mol/L，pH 9.0）

将 4.2 g 碳酸氢钠溶解于 40 ml 水中，用 5 mol/L NaOH 将 pH 调节至 9.0，并用水定容至 50 ml。

20×SSC

TE（pH 8.0）

酶和酶缓冲液

10×PCR 缓冲液

100 mmol/L Tris-HCl（pH 8.3）

500 mmol/L KCl

Sequenase 测序酶（13 单位/μl）（US Biochemical）

Sequenase 测序酶稀释缓冲液（US Biochemical）

5×Sequenase 测序酶反应缓冲液（US Biochemical）

Taq DNA 聚合酶（5 单位/μl）（如来自 Applied Biosystems 的 AmpliTaq Gold）

核酸和寡聚核苷酸

100×氨基-烯丙基标记的 dNTP 储存液

25 mmol/L dATP

25 mmol/L dGTP

25 mmol/L dCTP

10 mmol/L dTTP

15 mmol/L 氨基-烯丙基标记的 dUTP

461

对照 DNA，剪切的酵母基因组 DNA（5 μg/ml）

dATP、dCTP、dGTP、dTTP（100 mmol/L 储存液，超纯级）（如 27-2035-01，Amersham Bioscience）

免疫沉淀的 DNA，按照方案 10 中介绍进行提取

poly（A）/poly（dA）（5 μg/μl 多聚腺苷酸）（Sigma）

引物 A 寡聚核苷酸：5′-GTT TCC CAG TCA CGA TCN NNN NNN NN（40 μmol/L）

引物 B 寡聚核苷酸：5′-GTT TCC CAG TCA CGA TC（500 μmol/L）

酵母基因间隔区微阵列（见第 1 章，方案 2）

专用设备

微量移液枪用带气溶胶滤芯吸头

对于所有不使用多通道移液器的步骤，使用气溶胶滤芯吸头以避免污染。

扩增管（用于 PCR 的 0.2 ml 薄壁试管）

微阵列阅读/扫描仪

Microcon YM-30 微型浓缩装置（Millipore）

移液器（可调式，单通道）和带滤芯吸头

移液器（容量 2.5 μl 和 200 μl）（12 通道，容量可调）

QIAquick PCR 纯化试剂盒（QIAGEN）
热循环仪，可编程并带有加热顶盖（如 ABI 9700）

附加试剂和设备

本方案的步骤 12 需要第 3 章，方案 14 中列出的试剂。
本方案的步骤 18 需要第 3 章，方案 20 中列出的试剂和设备。

方法

A 轮：免疫沉淀 DNA 的扩增

1. 制备足够量的反应混合物，足够用于预计的扩增反应总数加上一个反应（见步骤 2）。在置于冰上的 0.5 ml 微量离心管内合并下列各组分。在反应混合物使用之前加入 Sequenase（在步骤 3 之中）：

5×Sequenase 测序酶反应缓冲液	1.0 μl
A 轮 dNTP 混合物（储存液中有 4 种 dNTP，每种的浓度均为 3 mmol/L）	1.5 μl
二硫苏糖醇（DTT）（0.1 mol/L 储存液）	0.75 μl
牛血清白蛋白（0.5 mg/ml 储存液）	1.5 μl
Sequenase（13 单位/μl）	0.3 μl

2. 为每个扩增反应，在各自的加盖的 0.2 ml 薄壁 PCR 管内合并下列试剂。

免疫沉淀的 DNA（或对照样品）	7 μl
5×Sequenase 反应缓冲液	2 μl
引物 A（40 μmol/L 储存液）	1 μl

   作为对照，建立包含以下组分的反应混合物：
   - 用剪切的酵母基因组 DNA（5 μg/ml）替代免疫沉淀的 DNA；该反应的扩增产物将用作参考探针（如果需要时）。
   - 无 DNA（用 7 μl TE 替代 DNA）

462

3. 将步骤 2 中准备的装有 DNA/引物混合物的试管放入带有加热顶盖的热循环仪中。使用下列的变性、复性和聚合时间与温度进行扩增：
   a. 94℃下 2 min
   b. 8℃下 2 min

      根据需要，本步骤之后需要编程或者输入一次暂停。
      在 8℃下暂停期间，向每支试管内加入 5 μl 包含 Sequenase 的反应混合物（来自步骤 1）。迅速从热循环仪中依次取出试管，加入反应混合物，用移液器上下吸吹液体数次，以确保混合均匀，并将试管重新放回仪器中。如果需要更多的时间来加液和混匀，可以在温度开始沿斜线上升至 37℃之前将仪器设置为“暂停”。

   c. 在 8 min 的时间内，使温度沿斜线上升至 37℃。

      并非所有的热循环仪都可以如此缓慢的进行升温。在 ABI 9700 上，升温速度为（正常值的）7%可以实现所希望的速度。

d. 将反应在 37℃下保持 8 min。

4. 重复步骤 3 中描述的扩增循环的 a～d，并且在步骤 3b 处加入额外的 Sequenase：
   a. 为每个反应准备包含 0.3 μl Sequenase 和 0.7 μl Sequenase 稀释缓冲液的反应混合物。
   b. 在第二个 8℃下的暂停期间（第二轮的步骤 3b），向每支试管内加入 1 μl 新鲜配制的 Sequenase。用移液器将酶加入每支试管内的 15 μl 液体中。
   c. 使用另一支设定为 8 μl 的移液器，上下吹打混匀试剂，并在下面的步骤（温度沿斜线上升至 37℃）开始之前将试管放回热循环仪中。
5. 将反应冷却至 4℃。
6. A 轮完成之后，用 TE 将产物稀释到 50 μl。使用 15 μl 稀释后的 A 轮产物用于下面的扩增步骤，并将剩下的产物在－20℃下保存。将每份 15 μl 的样品加入到 0.2 ml 的薄壁 PCR 管中。

B 轮：免疫沉淀 DNA 的扩增和氨基-烯丙基标记

7. 混合下列组分，为即将在 B 轮中扩增和标记的各反应作准备：

双蒸水	63.0 μl
10×PCR 缓冲液	10.0 μl
$MgCl_2$（25 mmol/L 储存液）	8.0 μl
引物 B（500 μmol/L 储存液）	2.5 μl
100×氨基-烯丙基-dNTP 储存液	1.0 μl
Taq DNA 聚合酶（5 单位/μl）	1.0 μl

8. 向各支装有 15 μl A 轮中扩增的 DNA 的 0.2 ml 薄壁 PCR 管内加入 85 μl 步骤 7 中准备的反应混合物。

463

9. 将试管放入热循环仪中，并使用下列的变性、复性和聚合时间与温度扩增 DNA 模板：

循环次数	变　性	复　性	聚　合
32	92℃，30 s	40℃，30 s，然后 50℃，30 s	72℃，1 min

对 5 μl B 轮产物进行琼脂糖凝胶电泳分析。应该看到拖尾的 DNA，大小为 200～600 bp。在含有“无 DNA”对照的样品的泳道中应该见不到扩增的 DNA。

10. 向 Microcon YM-20 装置中加入 400 μl H_2O。每个扩增混合物使用一个单独的装置。将每个 B 轮扩增反应浓缩至约 20 μl。向同一个 Microcon YM-20 装置中加入 500 μl H_2O 并将溶液浓缩至体积在 19 μl 以下。用水将回收样品的体积调整到 19 μl。避免干燥样品，否则随后它们难以重悬。用 9 μl 每种浓缩的氨基-烯丙基-标记样品与 Cy 染料耦联。将剩余的样品于－20℃下保存在 H_2O 中。
11. 向 9 μl H_2O 中的浓缩样品内加入 1 μl 新鲜配制的 1 mol/L 碳酸氢钠缓冲液（pH 9.0）并混匀。
12. 将 10 μl 溶于 0.1 mol/L 碳酸氢钠中的 DNA 样品加入一份分装好的合适的琥

珀酰亚胺（NHS）酯连接的 Cy 染料中。上下吸打混合试剂，确保经过分装的干燥 Cy 染料完全重悬。让染料与氨基-烯丙基标记 DNA 的耦联反应在室温下避光进行 1 h。

关于如何分装和保存单功能 Cy 染料，见第 3 章，方案 14。

13. 向每个耦联反应中分别加入 70 μl 水和 8 μl 3 mol/L 醋酸钠（pH 5.2）。从未偶联的染料中纯化标记的 DNA，每个耦联反应单独进行，按照下面的介绍使用 QIAquick PCR 纯化试剂盒：
    a. 向每个耦联反应中加入 700 μl PBS。将混合物上样到离心柱的上部。将样品在普通的离心机中以 10 000 g 离心 1 min。弃去流出液体。
    b. 向每个样品中加入 600 μl PE 缓冲液并与步骤 a 进行同样的离心。重复该步骤一次。直到最终的流出液体被弃去，将柱子以 13 800 g 离心 30 s 甩干。将离心柱转移到一个新的微量离心收集管内。
14. 向柱子的膜中央加入 60 μl EB 缓冲液（洗脱缓冲液）。室温下将柱子竖立放置 5 min，使 DNA 回收最大。以最高速度离心柱子 1 min，在普通的微量离心管中收集洗脱液。重复本步骤一次，在相同的试管内收集洗脱液。
15. 合并适当的 Cy3 和 Cy5 标记的纯化的探针，并用 TE 将体积调整到 500 μl。使用 Microcon YM-20，浓缩至约 10～15 μl。加入 H_2O、poly（A）/poly（dA）和 20× SSC，使体积满足微阵列的几何形状及盖玻片类型的要求，并且最终浓度为 3.5×SSC。
16. 加入 10% SDS 至最终浓度为 0.1%，煮沸探针，将溶液加到含有酵母基因间隔区单元的微阵列的表面。在 65℃下杂交 6 h，然后清洗并扫描微阵列。

关于杂交，清洗和扫描的详细情况，见第 3 章，方案 20 和表 3-9。

PART Ⅳ　SNP 序列：作图数据

464

Kerstin Lindblad-Toh（怀特黑德研究所/MIT 基因组研究中心，剑桥市，马塞诸塞州，02139）；
Thomas Ryder（昂飞公司，圣塔克拉拉，加利福尼亚　95051）

大规模基因分型方法得益于在最初的大规模 SNP 探索计划之一中鉴定出来的 SNP 标记（Wang et al. 1998）。基于 Lindblad-Tohet al.（2000）介绍的方法，方案 9 中描述了应用 Affymetrix HuSNP 阵列和 HuSNP 作图分析来检测杂合性丢失（LOH）的方法。

表 6-7　多重引物集合：SNP 作图数据

表中汇集了用于第一轮多重扩增反应的位点特异性引物。引物集合由 T3 和 T7 加尾的引物对组成，并且 24 个集合中的每一个代表了 50～100 个遗传位点的散在集合。

表 6-8　Affymetrix HuSNP 芯片中的 SNP 在图谱上的位置

表中给出了 1081 个 SNP 的作图位置。其中包含的许多 SNP 是和 Whitehead Institute/MIT Center for Genome Research（www-genome.wi.mit.edu/SNP/human/index.html）一道发现的。这些 SNP 是使用来自 Stanford Human Genome Center（www-shgc.Stanford.edu/RH/index.html）的详尽的 TNG 辐射杂种版进行作图的。

表 6-6 多重引物集合：SNP 作图数据

WIAF 命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi 编号	NwLib chrom
WIAF-1000	C/T	C	T	TCAAGATCCTCCTCAAACTT[C/T]AAGGGTGAAAAGCATACCAT	TAAAACATATCAAGATCCTCCTCA	AATGGTATGCTTTTCACCCTT	3099	15
WIAF-1001	C/T	C	T	CATTCTGTTCTCCAGTCTTG[C/T]AGGTTACATAAGCCATTTCC	CACATTCTGTTCTCCAGTCTTG	TTATGGAAATGGCTTATGTAACC	3100	18
WIAF-1002	A/T	A	T	GTGCTATTCAAGCAACAATT[A/T]TCTTTTATGTTCCTAAGCTC	GGGTGCTATTCAAGCAACAA	TCATGATGAGCTTAGGAACATAAA	3101	18
WIAF-1004	A/G	A	G	CCTTAAAATCACAGTACTGG[A/G]ATCAGGTGAAATTTGCTTTC	CGTTTCCTTAAAATCACAGTACTGG	TTAGAAAGCAAATTTCACCTGAT	3103	2
WIAF-1007	C/T	C	T	GCTACCCAGGCTGGTCTCAT[C/T]TCAGGCTCATGCGATCCTCC	CTACCCAGGCTGGTCTCAT	GGATCGCATGAGCCTGA	3106	3
WIAF-1008	G/A	A	G	GAAATCGACTACCAGCTGAT[G/A]AAATACCTGCAAAGTGCTCT	GAGAAATCGACTACCAGCTGA	TTTAGAGCACTTTGCAGGTATTT	3107	5
WIAF-1009	T/C	C	T	TCACTTCATGTGAAAACTAC[T/C]CCAGTGGCTGACTGAATTGC	TTCTTCTCACTTCATGTGAAAACT	CAATTCAGTCAGCCACTGG	3108	6
WIAF-1012	T/A	A	T	GTTCTGGTACCAATATCAAT[T/A]CAGTTTTCAAAGCCATTTGC	TGGGTTCTGGTACCAATATCAA	GCAAATGGCTTTGAAAACTG	3111	15
WIAF-1013	G/C	C	G	TCACAGTGCTGTGTCCTGCA[G/C]GCTGTCCTCAGGCAAGTTGC	ACAGTGCTGTGTCCTGCA	CAACTTGCCTGAGGACAGC	3112	1
WIAF-1014	T/C	C	T	GCCCTTTCAATTAATAGTAG[T/C]TGAGCGCTGGGGGCTGAAGT	TCAATGCCCTTTCAATTAATAGTA	AGCCCCCAGCGCTCA	3113	1
WIAF-1018	T/G	G	T	TAACCTCAATTCGTGAAAAC[T/G]AACATGCCTCAAAAAAGAGG	TCTAATAACCTCAATTCGTGAAAA	CAGCACAGTGTTTCTGTTAAAGTTT	3117	1
WIAF-1021	C/G	C	G	TGAATCATAAATTATCAAAT[C/G]TGCGCCCAGTAATCTGTGTT	CCCACTGAATCATAAATTATCAAA	CACAGATTACTGGGCGCA	3120	17
WIAF-1023	A/G	A	G	GAGGCATTAACTGAGATTAT[A/G]GGAAACGCACAGCAAAATTG	CCCAGAGGCATTAACTGAGATT	ATTTTGCTGTGCGTTTCC	3122	2
WIAF-1024	A/G	A	G	TCACCGAACAATATTTTAGG[A/G]ATTTGAAATTATTTCTGTAC	CTTTCACCGAACAATATTTTAGG	CATATTCTTGGGTGGTGAGAA	3123	3
WIAF-1026	T/C	C	T	TGAGTAATGCATTTGAGCTG[T/C]CCCAGGCTCTGTCTCCTCAG	CAGTGAGTAATGCATTTGAGCT	GAGGAGACAGAGCCTGGG	3125	14
WIAF-1027	A/T	A	T	CGAAGATTTTGGCTAATTTG[A/T]TAGTCTTACAAAGGCAGTCT	TCGAAGATTTTGGCTAATTTG	GGTGAACTAGACTGCCTTTGTAAGA	3126	8
WIAF-1029	A/G	A	G	CCACCTGTCTTCTCTCCCAC[A/G]TGCACAGCTTCCTGAGTCAC	CCACCTGTCTTCTCTCCCA	TGACTCAGGAAGCTGTGCA	3128	8
WIAF-1031	G/C	C	G	AACACCAGCGTCCCCCACCC[G/C]CGTCGTGGTGTAGTCATAGA	CAGCGTCCCCCACCC	GCTCTATGACTACACCACGACG	3130	2
WIAF-1032	C/T	C	T	CGATTTAATACGTACTTTAG[C/T]TGGAATTATTCTATGTCAAT	CCAGGCGATTTAATACGTACTTT	TTTCATAGCTTAAAAATCATTGACA	3131	4
WIAF-1036	A/G	A	G	TTCTATTCCCACGTTAGCCA[A/G]TTGTTCTTGATGAATCTATA	TTCTATTCCCACGTTAGCCA	TGACTCATATAGATTCATCAAGAACA	3135	14
WIAF-1040	T/C	C	T	AGGTGCTACCAGAACTCATG[T/C]GATAGCGCTTTCTTTTAGGC	TAAGGTGCTACCAGAACTCATG	TGCCTAAAAGAAAGCGCTATC	3139	13
WIAF-1044	C/A	A	C	ATGATGCCGAGGTCACTGCT[C/A]TACAAAGATTAAAGAAACTT	TGCCGAGGTCACTGCT	GGAAGTGTTTGATGGTAAGTTTCT	3143	8
WIAF-1045	A/G	A	G	TGATACACTACGCATGCACA[A/G]ATAAAGTCACATCAAGACTA	TGATACACTACGCATGCACA	CATTTAGATTATTAGTCTTGATGTGACTT	3144	3
WIAF-1046	T/C	C	T	CATTGGTCTAGGAGTAGAAA[T/C]GCACACAAGGAATAAGGGAG	GCCATTGGTCTAGGAGTAGAAA	CTCCCTTATTCCTTGTGTGC	3145	5
WIAF-1049	C/T	C	T	AGCTAAATTAGCCCTAAATG[C/T]GGGTAATATTTTTCCTCATG	TGACAGCTAAATTAGCCCTAAATG	TTTTAAAACATGAGGAAAAATATTACC	3148	5
WIAF-1051	A/G	A	G	GTCTCCCTGAGGTGATATGG[A/G]CCTTAAGTCCACGATGGTGA	TGTCTCCCTGAGGTGATATGG	TCACCATCGTGGACTTAAGG	3150	9
WIAF-1053	C/T	C	T	ATTGCGAAATCTGTTGTGCA[C/T]GCTCAAATGAAAACGCCTTT	TGCGAAATCTGTTGTGCA	AAGGCGTTTTCATTTGAGC	3152	12
WIAF-1056	T/G	G	T	GTAGATAAGAGGCATAATCT[T/G]AAACAAAATTCTTTCTGAAA	CTCTCAAGTAGATAAGAGGCATAATCT	GCTAAATTTTCAGAAAGAATTTTGTTT	3155	10
WIAF-1057	G/C	C	G	TGGAACCAACGATCATATCT[G/C]TATGCCTCATTTTATTCTGT	AGTGGAACCAACGATCATATCT	CATTGACAGAATAAAATGAGGCA	3156	13
WIAF-1059	C/T	C	T	TTGTTCCTCCAGACCGCAGG[C/T]TCCCCCAGCCTCAGGTTGCT	TCCTCCAGACCGCAGG	ACCTGAGGCTGGGGG	3158	11
WIAF-1063	C/T	C	T	ACATGAATACTTGAAGACAA[C/T]ATAACTACAACCTTACAAAT	CTACATGAATACTTGAAGACA	GGCATTTGTAAGGTTGTAGTT	859	13
WIAF-1082	C/T	C	T	AAAATTTTCCAGTTGGTAAG[C/T]AGCAGGTGCCGAGGGTCTGG	GCATAAAATTTTCCAGTTGGTAA	ACCCTCGGCACCTGC	19381	19
WIAF-1084	C/T	C	T	ATATTGCCAGCTTTGTAAAG[C/T]CATTAAAGCAGAAGTTTCTT	GGATTGTAATATTGCCAGCTTTG	AAGATCACTGAAGAAACTTCTGCTT	19368	6
WIAF-1085	A/G	A	G	TGAATGCTCTGCCATGAGCC[A/G]CAAGCAGCACAGTGATCATC	CACTGAATGCTCTGCCATGA	AACCTGTCCTTGTGGGTGAT	19373	8
WIAF-1089	C/T	C	T	GCTTCCTTGACTGTTCTCCA[C/T]GCCAGATTCTTATCAATGAT	TGCTTCCTTGACTGTTCTCCA	TGAAAGATCATTGATAAGAATCTGG	19480	8
WIAF-1091	A/C	A	C	ACCGCGTATAGTAAGCTCTG[A/C]GGAACTCCAAGAATCTAGAG	AAGATTAGGACAGACCGCGT	GATCTAAGCAGCGTTCCCAC	19385	14
WIAF-1119	C/T	C	T	GGAGTTTTGTTCTTTGTAAC[C/T]CTCTCATCATCGAGGCTATA	TTTGGAGGTTTGTGTCTGGA	TTAATATATAGCCTCGATGATGAGAG	19435	18
WIAF-1121	C/T	C	T	CATCTGGGAAGATTTCTACC[C/T]GACCAACAGTTCCTTCAGCT	TGCATCTGGGAAGATTTCTACC	GGAAGCTGAAGGAACTGTTGG	19361	1
WIAF-1133	A/G	A	G	CAGCAGCAGGTGGCAGGAAG[A/G]AGTGGAGGGAAAGGACACCA	GTGCCCTGAAGATTAGCAGC	ACTTGGTGTCCTTTCCCTCC	19376	17
WIAF-1134	A/G	A	G	AAGTGGGCCCTCCCAGTCCC[A/G]TCTCTGGGCACAGATCCCAC	AAGTGGGCCCTCCCAGTC	GAGCAGACTGGTGGGATCTG	19436	17
WIAF-1138	A/G	A	G	TCCACAGCGCAATGACAGCA[A/G]CCTCTCTCCCACCCACTCAA	AGTGATTTTGGCTAGGCGTG	CTTGAGTGGGTGGGAGAGAG	19481	17
WIAF-1139	A/G	A	G	TAATTTTGCAGGCAGACTTC[A/G]TAGAGCCATTCTGTGCAGAA	CGTAATTTTGCAGGCAGACTT	TCTTCTGCACAGAATGGCTC	892	16
WIAF-1145	C/T	C	T	TTTGTGAACGTCACATTGGC[C/T]GAGCCGTGTTCAGTTCCCAG	TTTGTGAACGTCACATTGGC	TGGGAACTGAACACGGCT	19437	10
WIAF-1146	C/T	C	T	AACTAAGTGTTATTGAACTG[C/T]AATTCTCTCAATAACTCAGT	TGAAGCCAACTAAGTGTTATTGAA	GTGCTACACTGAGTTATTGAGAGAA	897	2
WIAF-1148	A/G	A	G	CAAAATCAAATATTCATCAC[A/G]TTGGGTTGAAAAGTTGGAAG	AACAATTCACAGCTACAGGAAATC	TCTTCCAACTTTTCAACCCAA	19438	19
WIAF-1149	C/T	C	T	GAAACCAGAGTTTCCCCAGG[C/T]GAGCCCTTCCTATCTGCGGT	ATGACTTTCCTGTCCCATCG	TAACCGCAGATAGGAAGGGC	19186	9
WIAF-1154	C/G	C	G	GATGATGGGGTCTGAGATGT[C/G]TCCTCAGGCTGCATCAGCTG	TTTCAGATGATGGGGTCTGA	TGGGTCAGGCTCTTTCTGTT	19414	1
WIAF-1156	A/G	A	G	CTATACCAAGTTTCCATAAA[A/G]CTGTCTGCTGGTTGGGGAGG	TTTAGCTGCTATACCAAGTTTCCA	GCAAAGAATGTGGTCAGGCT	19378	16
WIAF-1162	A/G	A	G	TCTTGGCATTCGTCCTCCTC[A/G]GGCCAGTGCTCCACCCAAGT	CTGGTTCTCTTCGTCTTGGC	ATCATCGGGAAGGACACTTG	19439	19
WIAF-1164	A/T	A	T	TCAGTTCACAGTGGAGACTA[A/T]GGAGATTCAGGGCAGGATCC	CTCCCCTCCTCAGTTCACAG	GGATCCTGCCCTGAATCTC	19365	17
WIAF-1165	C/T	C	T	TTCTCCTCTTCCAGAAGTTG[C/T]AGACGTCTATTTAGTTTGAT	GCACGTTCTTTGTTCTCCTCTT	CGACAGATAATCAAACTAAATAGACG	19371	

续表

WIAF 命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi 编号	NwLib chrom
WIAF-118	C/T	C	T	CTATTGTGGAGAACAAACAG[C/T]TTGGGAAGTAAAGGTTGATT	GCCTATTGTGGAGAACAAACAG	AAGTAATCAACCTTTACTTCCCAA	2512	19
WIAF-1186	A/G	A	G	TGCAGATCTTCTATTCCTGG[A/G]GTTGATATGACAAGGAAACC	CACTGCAGATCTTCTATTCCTGG	CCAATAGGGTTTCCTTGTCA	742	8
WIAF-1218	C/T	C	T	GTTTGCTGGGACCTGGAACA[C/T]TGGACTTCTTTCTACTGCAG	TGCTGGGACCTGGAACA	CTGCTGCAGTAGAAAGAAGTCC	19482	7
WIAF-1242	A/T	A	T	AAGCCAGTATGTTGTGGCAA[A/T]ATTCGAGAAAACACACTGAA	GAAGCCAGTATGTTGTGGCA	TTTTTCAGTGTGTTTTCTCGAAT	19483	15
WIAF-1247	A/G	A	G	GAAGGGCAGGACTAATCCAA[A/G]TCTCTACCCGCAGCTTGCTC	GAAGGGCAGGACTAATCCAA	GAGCAAGCTGCGGGTAGAG	19484	11
WIAF-1250	A/T	A	T	ATCTTCAGGAAAATGAAACA[A/T]GAGAAGCTGAAACAATCTAC	GCCTGCTCACAAGGTAGACA	CATTCAGGTGTAGATTGTTTCAGC	19380	5
WIAF-1251	A/T	A	T	AAATTACTGCCTTAGCAAAC[A/T]CTATGTTGTCAGGTTTTTCT	GGTAAAGTCTAAATTACTGCCTTAGCA	TGCAGCAGAAAAACCTGACA	19440	5
WIAF-1252	C/T	C	T	CAGTTTGGCTTCTGTCCTCA[C/T]AGTCTCTCTCCATGTGGCAA	CCAGTTTGGCTTCTGTCCTC	TGTTTGCCACATGGAGAGAG	19362	1
WIAF-1258	C/T	C	T	CTTAAGTACTGTATCGCTGT[C/T]ATGCAGCGGCCTGTGGAGGC	TCTTTGGTTGTCTACACAGACACTT	CTCAGGACACAGGCCCAG	19417	1
WIAF-1262	A/G	A	G	TACATGGCTAACCTTCTAAC[A/G]CTTGAAAAATCAATTTCAAG	CCTTTACATGGCTAACCTTCTAAC	TCCCTTGAAATTGATTTTTCAA	929	10
WIAF-1266	C/G	C	G	ACAGGAGAGCGCAAAACAAA[C/G]CCTGGCTGCCTCGGGATGGA	GCAGCATTTGTGACAGGAGA	ATGCAGTGGTGGTGAGGC	19441	22
WIAF-1274	C/T	C	T	CTCTGGTGTGCAGGGAATCA[C/T]TTTGCTGGATTAGAGGAAAG	GTCTCTGGTGTGCAGGGAAT	AAGTCATGGAAACAGACGGC	19375	17
WIAF-1283	A/G	A	G	TGATTGCCTCACACAAGCAT[A/G]ATCAATCGCCACGAGAGACT	ACCATGATTGCCTCACACAA	CCAGCCATACTCTTTGGCAT	19374	15
WIAF-1284	C/G	C	G	AATCCAGTATATGCAGAAAG[C/G]TACTCAGCATCACACTCGTG	AAAGTACAATCCAGTATATGCAGAAAG	TGATCACGAGTGTGATGCTG	19366	17
WIAF-1290	C/T	C	T	ACCTATCTGAGTCTGTATCC[C/T]CATCTTTAAATTGTAAATTT	ATCTCACCTATCTGAGTCTGTATCC	TGATAGGTGTCCTTAAAATTTACAATTT	19431	X
WIAF-1291	C/G	C	G	TGTGTGGCAAAATAGAAAAT[C/G]TTTCAATTACATTAGGAAAT	CCAAGTGTGTGGCAAAATAGAA	CCCGATTTCCTAATGTAATTGAA	934	8
WIAF-1302	A/G	A	G	TTCAGCCTCCTCTACCCTAC[A/G]AGATCTGGAGCAACAGCTAG	CCCTGTAGCAGTCTTCAGCC	TTTCCTAGCTGTTGCTCCAGA	19442	4
WIAF-1306	A/G	A	G	ATAGCTGTTCCTGAGTTGCA[A/G]GCACGATGGAGATTTGGACA	AGATAGCTGTTCCTGAGTTGCA	CAGTGTCCAAATCTCCATCG	19443	6
WIAF-1307	A/G	A	G	CTCTGCCTGCAGGATGTGCG[A/G]CGTGTGCCTGGAGTAGCCCC	TCTCTGCCTGCAGGATGTG	GGGCTACTCCAGGCACAC	19159	11
WIAF-1310	C/T	C	T	TTCCTCCACTTCATTTTTCC[C/T]TTGTCCTTTCTCTCTAATTC	TCTTCCTCCACTTCATTTTTCC	AAACACTGAATTAGAGAGAAAGGACA	946	1
WIAF-1318	C/T	C	T	AGCCATTTCTTTTTGAATGA[C/T]GACAATAAAAATGATTTGAA	CAGCCATTTCTTTTTGAATGA	TTCCCAATTTCAAATCATTTTT	753	13
WIAF-1320	C/T	C	T	TTTATTTCTTTCTTTCCATT[C/T]TATGTTGGTAAATATAAAGA	TGTGTTACTTTTATTTCTTTCTTTCCA	TTGCACAATCATCTTTATATTTACCAA	755	8
WIAF-1325	C/T	C	T	CTGCAGTACAAGGATCTGCC[C/T]GTCTCCCTGCGTCGGGGGGC	CTGCAGTACAAGGATCTGCC	CCCGACGCAGGGAGA	19433	17
WIAF-1331	G/C	C	G	GTATTGAAGTCTTTGGGCCA[G/C]GTCTGTAATGATCTGACTTC	GTATTGAAGTCTTTGGGCCA	GGGAGAAGTCAGATCATTACAGA	2285	5
WIAF-1353	C/G	C	G	TGACTAGCCATCTCAAGCAA[C/G]TTTCGACGTTTGACTAGCCA	TGACTAGCCATCTCAAGC	TGGCTAGTCAAACGTC	962	13
WIAF-1362	C/T	C	T	TGGTCATCGTGGCCATCGCC[C/T]GGACTCCGAGACTCCAGACC	ATCGTGGCCATCGCC	TGGTCTGGAGTCTCGGAGTC	19434	8
WIAF-1366	A/G	A	G	TTTTCAGTATGGTTCAGAAT[A/G]AGTATCTTAGTATTCTTTCT	GCTGCCACTTTTCAGTATGG	CCATAGCAAAATAGAAAGAATACTAAGA	965	6
WIAF-1412	A/G	A	G	CTGGTTTTCTAGATGTACAC[A/G]TGTGGGACCTCTGTCTCAAC	CACCCAACTGGTTTTCTAGATG	AGGTTGAGACAGAGGTCCCA	777	9
WIAF-1417	C/T	C	T	TTAGCAAACATTTTTTAAAA[C/T]CCACATCCAACAGATTGGTT	TCTTACTTGGGTAGTTTAGCAAACATT	AACCAATCTGTTGGATGTGG	19372	1
WIAF-1421	A/T	A	T	CTGATGAGGAAGCTCTAGTG[A/T]AGAAATTCAGGACGCGGTCT	GGGTAACCTGATGAGGAAGC	GACTTGAACCAAGCCCTCTG	19444	17
WIAF-1430	A/G	A	G	CACAGCAGCTTACTCCAGAG[A/G]TCAAGTCCAAGGCCATTGGC	ACACAGCAGCTTACTCCAGAG	CCAATGGCCTTGGACTTG	674	12
WIAF-1450	C/T	C	T	AAGCAGCCCTCTGCAAGTGG[C/T]GCTGGATACCACTAAGAAGT	CAGCCCTCTGCAAGTGG	TGCAGTAGACTTCTTAGTGGTATCCA	993	9
WIAF-1461	T/A	A	T	CCATAGGACAGTTGTTTGAG[T/A]TTGGTGCCACCTTATTGGCC	ATTCCATAGGACAGTTGTTTGAG	TCTACACCCCAGTTTCCTTATC	2004	6
WIAF-1463	A/T	A	T	CCAATGAAGCAGGCTATGGA[A/T]TAAAAGAAACAACAAATGGG	CCAATGAAGCAGGCTATGG	AATGAACCCATTTGTTGTTTCT	1090	11
WIAF-1464	T/C	C	T	GAATGCCTACATCTGGAATT[T/C]CATTACATCAACGTTAAATT	GGGAATGCCTACATCTGGAA	GGACAAAATTTAACGTTGATGTAATG	3161	
WIAF-1465	G/T	G	T	GAAACTGCATGGGTACAAAT[G/T]TCCAATTCATACTTAACAAG	GGAAACTGCATGGGTACAAA	TCCCCACCTTGTTAAGTATGAA	3162	11
WIAF-1466	A/G	A	G	GAACTTGATGGAAAATTGGC[A/G]TCTGTGTAGAATGATTCTAA	GAACTTGATGGAAAATTGGC	AAAGCTTTAGAATCATTCTACACAGA	3163	4
WIAF-1468	G/A	A	G	AGGCCTGAGGTGGGGCTTAC[G/A]AATTAGTATTTCGAAGACCT	GCCTGAGGTGGGGCTTAC	TTAGGAAGGTCTTCGAAATACTAATT	3164	18
WIAF-1469	A/G	A	G	TAAACACACATAAAGACACC[A/G]GGTTCTCAGTAATGCTCTAG	ACACATTCATTTTTATAAACACACATAA	CCTGGACTAGAGCATTACTGAGA	3165	4
WIAF-1470	A/G	A	G	TGGCTGTGCCAACTCAAGGC[A/G]GCTACCTTGACATTATAAGG	GCTGTGCCAACTCAAGGC	CAGCCTTATAATGTCAAGGTAGC	2006	1
WIAF-1473	G/A	A	G	GTCAAAAACCACATGTTCTC[G/A]TAAGTGGGAGATAAACAATG	TAATGATACAGAAAGTCAAAAACCA	CACGTCCAGGTGTACACATTG	1092	3
WIAF-1474	A/G	A	G	AATTAGTTTTCTGTAATATT[A/G]TTCTAGTCCATTTAGATTGT	CCAACCTAACAAATTAGTTTTCTGT	CATTTACACAATCTAAATGGACTAGAA	3166	16
WIAF-1475	A/G	A	G	CAACTACATTAATTTTCTTG[A/G]TGAACACAGTAGACATATGA	GCACCCAACTACATTAATTTTCTT	AGTACATTTCATATGTCTACTGTGTTCA	1093	4
WIAF-1476	T/C	C	T	CACCTGTTTGTATAAACTAT[T/C]GTGGGGTAAGCCCTTCTTTG	CCAATGCACCTGTTTGTATAAA	AAAGAAGGGCTTACCCCAC	3167	6
WIAF-1477	C/T	C	T	ACCTCCTGCCTGCCTGCTCA[C/T]GACTGATTACTTTCAATCCC	CCTGCCTGCCTGCTCA	GCTGGGATTGAAAGTAATCAGTC	3168	7
WIAF-1479	C/T	C	T	TAAAATGACCTTGATGCCGT[C/T]GGAATTATATTCAGACACAG	AAAATGACCTTGATGCCGT	AAGCCTGTGTCTGAATATAATTCC	1094	17
WIAF-1482	A/T	A	T	AATGTCTTAACATACCAAAG[A/T]AGTGGAATCAATAGAATAAA	TTTAAAGAAATGTCTTAACATACCAAAG	CGTACGCTTTGTAAGACTTAAATATTTT	3171	8
WIAF-1488	G/A	A	G	CACAGGTAGTCCACAGAATA[G/A]GACACAAGAAACCTCAAGCT	GAGCACAGGTAGTCCACAGAA	CAGCTTGAGGTTTCTTGTGTC	3174	19
WIAF-1489	G/A	A	G	ATTCATTCTTTTCTTCACCA[G/A]TAAAGGCTGTTCTTGGGCTT	CGCCTACAAAGAATCATTCATTC	AGCCCAAGAACAGCCTTTA	826	3
WIAF-1490	T/C	C	T	GGTCTGGGCTTCTGAGGTCC[T/C]GGTAGAAGGAGGGCAGGTGG	TCTGGGCTTCTGAGGTCC	CACCTGCCCTCCTTCTACC	3175	15
WIAF-1500	C/T	C	T	ACCTGACCTCTGACATGTGG[C/T]CTCTGGTCCCCATTTGTCTC	CCTGACCTCTGACATGTGG	AGACAAATGGGGACCAGAG	3184	21
WIAF-1504	G/A	A	G	GCTGTGGTGGGGGGCAGTAT[G/A]AGCCAGGGACTCCCTTCCCA	GGAGCTGTGGTGGGGG	GAAGGGAGTCCCTGGCT	3188	1

续表

WIAF 命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi 编号	NwLib chrom
WIAF-1505	T/A	A	T	ATTATTAAAAGAGCAAAGTT[T/A]CCCCTCCCTTTCTTACTTTC	CTGAACATAATTATTAAAAGAGCAAAGT	TGAAAGTAAGAAAGGGAGGGG	2011	18
WIAF-1506	T/C	C	T	TTTGGACAACTTAAAACTTA[T/C]TAGTGACATTGCTGTCTAAT	CAATTAACTTTGGACAACTTAAAACTT	TTGATTATTAGACAGCAATGTCACT	2012	18
WIAF-1509	C/T	C	T	TCCTCGTCAAAAGAAAACCA[C/T]CCCGGGACCCACTGTTCTAT	TCCTCGTCAAAAGAAAACCA	ACAGTGGGTCCCGGG	2015	15
WIAF-1511	G/C	C	G	GCTTGAATTTTAGATGCAGA[G/C]ATCATTTTATATATATCTGT	CTGCTTGAATTTTAGATGCAGA	TGTAATTTTTCTGTATTACATTGACAGA	2016	10
WIAF-1514	G/A	A	G	GGCAGGGAGGGGCAGAGAAC[G/A]CACTAGCTTGGGGGTGGGCA	AGGGAGGGGCAGAGAAC	CCACCCCCAAGCTAGTG	3191	16
WIAF-1516	C/T	C	T	TCCCTCTAAAGTTTCAGTTT[C/T]TTCACCAGTAAAGGAAAAGG	GCTCATTCCCTCTAAAGTTTCA	CAACCTTTTCCTTTACTGGTGAA	3192	8
WIAF-1517	C/A	A	C	AAGATGCCCCATTCCTTGCG[C/A]CTCAGCCCTCAGTTCCTTCA	GCCCCATTCCTTGCG	GAAGGAACTGAGGGCTGAG	1100	8
WIAF-1519	G/A	A	G	GAGGGTACTGGAAGTGGGAG[G/A]AGGCAAGGTCTGCTATCAGG	GAGGGTACTGGAAGTGGGAG	CCTGATAGCAGACCTTGCCT	2018	17
WIAF-1520	A/G	A	G	ATCTTTTACAAAATTTCATT[A/G]GAACTGACAATGTTATGGTG	TGCCATCTTTTACAAAATTTCA	GACTCACCATAACATTGTCAGTTC	1101	12
WIAF-1524	C/T	C	T	TTTCCCCAACTACGATAAGT[C/T]CATTGCCGGATGCTGAATAA	CGTTTCCCCAACTACGATAAG	TCAGCATCCGGCAATG	2019	4
WIAF-1525	G/A	A	G	GGGATGACTGTAGAGGATTA[G/A]ATGATCCCTGAATACTTTTC	CAAAGGGATGACTGTAGAGGATT	GCTGAAAAGTATTCAGGGATCA	1103	11
WIAF-1529	A/G	A	G	TCATATCCTTTCCCTGCTAC[A/G]AAAGACAAAACAAAATGATG	TGTCATATCCTTTCCCTGCT	GGGCATCATTTTGTTTTGTC	2022	11
WIAF-1533	T/A	A	T	GCCAATGTCTTTTCCCAGTC[T/A]GTCATATTTTTGTCCATGGT	GCCAATGTCTTTTCCCAGTC	ACAACCATGGACAAAAATATGAC	1106	19
WIAF-1538	A/G	A	G	TCTTTTTACAATGCAGTTTC[A/G]ACATAACATTGGTAGAGTAA	TCTTCTCTTTTTACAATGCAGTTTC	GGTTTGTTGTTTACTCTACCAATGTT	3201	21
WIAF-1540	T/C	C	T	TTTACAAACCCATTTAAGTA[T/C]GGAATGATTTATATGGGGTC	GGTTTAGTTTACAAACCCATTTAAGT	CTGACCCCATATAAATCATTCC	1108	17
WIAF-1541	C/G	C	G	AAAGGTACAGCAGGCTCTAC[C/G]CATTCTCTCTCTCTTGCTCT	TTTACCAACTCGAAGAATTTAAAAG	TGTGTTGTGAGAGCAAGAGAGA	1109	6
WIAF-1542	T/A	A	T	CTTGTTTCATAAATAAGGGA[T/A]TTCAATCAAGATCCATGAAA	CAAGCTTGTTTCATAAATAAGGGA	CATTTCATGGATCTTGATTGAA	3202	11
WIAF-1543	G/T	G	T	AACATTAATCATTACTATTT[G/T]CAACTCCGTATTTTCAAAAC	AAGCCTTAAAAATAAAGCAACATTA	GTTTAGTTTTGAAAATACGGAGTTG	1110	1
WIAF-1544	G/C	C	G	TGGCTAGTCAAACGTCGAAA[G/C]TTGCTTGAGATGGCTAGTCA	TGGCTAGTCAAACGTCGAAA	TGACTAGCCATCTCAAGCAA	1111	13
WIAF-1549	C/T	C	T	AGAACCAGGGAGACACCTCA[C/T]GTGAGATGACCTGCACACAA	GAACCAGGGAGACACCTCA	TTGTGTGCAGGTCATCTCAC	1114	19
WIAF-1556	G/C	C	G	ACCCTCCCCATTTTGAGTCT[G/C]CATAGTCCATTATACCACTC	CACCCTCCCCATTTTGA	TGCAAAGGCATACAGAGTGG	3205	
WIAF-1559	C/T	C	T	TTCTTTCCCACTGAGCCTGG[C/T]TGAACTACAGCTGCCAGCAT	TTTCCCACTGAGCCTGG	TGCTGGCAGCTGTAGTTCA	3208	2
WIAF-1564	C/G	C	G	GGTTCTCCCCAGAGTAAAAG[C/G]GTTTTCTCTGGGACTTGGCA	GGGTTCTCCCCAGAGTAAAA	GCCAAGTCCCAGAGAAAAC	1122	11
WIAF-1565	A/G	A	G	CCCTGCCAACGTGCTGATCC[A/G]GTGCCTGTCCTGGAGTTACC	TGCCAACGTGCTGATCC	GGTAACTCCAGGACAGGCAC	2024	8
WIAF-1567	G/A	A	G	AAAGAAAATACTTGGTTAAA[G/A]TTGAAAGGACCTAGTTAGAG	GCAGACAAAGAAAATACTTGGTTAAA	TCTAACCTCTAACTAGGTCCTTTCAA	1124	6
WIAF-1577	A/G	A	G	GCCACGAGGAGTAGCCAGGT[A/G]GGCGAGACACAAAAAGGCCT	CCACGAGGAGTAGCCAGGT	GCCTTTTTGTGTCTCGCC	1130	1
WIAF-1580	A/G	A	G	TTGCTCATTTGCTACATGAA[A/G]CAGAGGCAGAGTATTCTGTG	TTTGCTCATTTGCTACATGAA	CCACACAGAATACTCTGCCTCT	1133	8
WIAF-1582	G/C	C	G	ACAGTTTTTATCCTGAAATT[G/C]TTTTTAGAAGCCCTGCCCAA	GGAGACAGTTTTTATCCTGAAATT	GGGCAGGGCTTCTAAAAA	2027	9
WIAF-1584	G/A	A	G	TCTGTAATGCCATTTTGATG[G/A]TTAACATTACGTGTCAACTT	GGTCTGTAATGCCATTTTGATG	GGCTTCCTTCATTCCATTCA	2029	8
WIAF-1585	A/G	A	G	GGGCAGAGTGAGGGGCAGAC[A/G]GGATGAGGCTCTTCTGTAAA	GCAGAGTGAGGGGCAGAC	GACTTTACAGAAGAGCCTCATCC	3213	10
WIAF-1588	G/A	A	G	GGAAAGCTGAGGCAGCTAAT[G/A]GGCTCATACAAAGGTTTGGA	GGAAAGCTGAGGCAGCTAAT	TCCAAACCTTTGTATGAGCC	3214	11
WIAF-1593	G/A	A	G	GCACTGGTGCACTTACAGGC[G/A]GAAGAGCTTCCTCATTTGCT	CACTGGTGCACTTACAGGC	CAGCAAATGAGGAAGCTCTTC	3217	4
WIAF-1594	G/C	C	G	CTTAATGTCTTGAGAACATG[G/C]AAATGAATTGGACAAAAATC	CATTCCTTAATGTCTTGAGAACATG	CAGTGATTTTTGTCCAATTCATT	2033	5
WIAF-1596	T/G	G	T	TGGGTCTCTTTCAAGTGAAT[T/G]TTCCTTTCGTTCCTGTTCTA	CTTGGGTCTCTTTCAAGTGAA	TTTAGAACAGGAACGAAAGGAA	3218	8
WIAF-1598	C/T	C	T	AACTCCCCAGGTTAAACACC[C/T]TGTAATTTACCTTGAAGCTC	AACTCCCCAGGTTAAACACC	TGTAAAGAGCTTCAAGGTAAATTACA	2035	17
WIAF-1602	T/C	C	T	ACTGAGTTACAGCCCCCTCC[T/C]CACTCCTGTATCTGCTCAGG	TGAGTTACAGCCCCCTCC	TCACCACCCTGAGCAGATAC	1139	14
WIAF-1606	T/C	C	T	TCTGCCCACCCTTTGAGCTG[T/C]ACCTGCCCCACCCCAGCTCA	GCCCACCCTTTGAGCTG	TGGGGTGGGGCAGGT	1143	22
WIAF-1607	G/C	C	G	AGGCTTGCCCTTCCCTGAAT[G/C]TGCCTTCTTGTGGCTTTAGC	GCTTGCCCTTCCCTGAAT	GCTAAAGCCACAAGAAGGCA	1144	22
WIAF-1610	G/A	A	G	ATAATGAAGTGAATAAACGC[G/A]TGTGAACTAATGTTTAAAAA	AAGCTTAATGAACATAATGAAGTGAAT	GCTCTAACTTTTTAAACATTAGTTCACA	3222	1
WIAF-1614	G/A	A	G	TCTCTCACGCTGATCCAGCC[G/A]GGCACCCTGCTTAAGTCAGT	TCACGCTGATCCAGCC	TGACTTAAGCAGGGTGCC	1146	16
WIAF-1620	T/C	C	T	ACAGGCCTGCATTTGGCTTA[T/C]GTGCCTGAAAAAGAAGGGCC	GGCCTGCATTTGGCTTA	GCCCTTCTTTTTCAGGCAC	2039	7
WIAF-1621	C/A	A	C	CAAAGCAGTTTTACTGTACA[C/A]AGAAGTGCAATGCTACATTA	AGTAAAACAAAGCAGTTTTACTGTACA	GGACTTAATGTAGCATTGCACTTC	2040	20
WIAF-1622	G/A	A	G	CTTCCTGGGCATGGCCGTGA[G/A]CACCCTGTGTGGCGAGGTGC	CTGGGCATGGCCGTG	CTCGCCACACAGGGTG	1149	8
WIAF-1623	A/G	A	G	TATTACATTCAGAATTATGC[A/G]AAACAATTAGTTATATTGCA	GCTGATTATTCATATCTATTACATTCAGAA	CAGCTTTGCAATATAACTAATTGTTT	1150	6
WIAF-1626	T/C	C	T	GTCATGGGAGATTGGATAGA[T/C]GCCTAACCTATCTCAATTTT	GGTCATGGGAGATTGGATAGA	CATTACTTAAAATTGAGATAGGTTAGGC	3226	18
WIAF-1627	T/G	G	T	GCATCTATACCTTCGAAACC[T/G]CCTCTTAACCTCTCCCAGGC	AATAGCATCTATACCTTCGAAACC	GCCTGGGAGAGGTTAAGAGG	3227	7
WIAF-1628	C/G	C	G	GGCAAGAGAGGCATCACTGA[C/G]GCTGCATCTGCCATGCGCTC	GCAAGAGAGGCATCACTGA	GCATGGCAGATGCAGC	1151	16
WIAF-1629	A/G	A	G	AAAGGAGGCAACTCTTTAAT[A/G]AAGGAAATGTTACCAAATCC	AAGATAAAGGAGGCAACTCTTTAA	CTATGGATTTGGTAACATTTCCTT	3228	3
WIAF-1634	T/C	C	T	ATCTTCAGCACAGGAATTGT[T/C]AATCTTCTCACTTCACCTGT	CATTATCTTCAGCACAGGAATTG	AAGAACAGGTGAAGTGAGAAGATT	2043	9
WIAF-1635	C/T	C	T	CTACGCAATTTTCATGCAGA[C/T]TGTGCACACAGTACAGTGCA	CTACGCAATTTTCATGCAGA	TGCACTGTACTGTGTGCACA	3232	15
WIAF-1638	C/T	C	T	AGACATTAGAAAAGGACGCT[C/T]CTGTTGGCTGAACAAAGATC	GGAAGACATTAGAAAAGGACGC	TGATCTTTGTTCAGCCAACAG	1154	9
WIAF-1642	A/T	A	T	TGACCCAATATCATTGAGGC[A/T]AACAGTTTGGGCTGTTTTTC	TGACCCAATATCATTGAGGC	GGAAAAACAGCCCAAACTGT	1157	1
WIAF-1643	A/G	A	G	AGTATCAACTTGAGTACCTC[A/G]TTATGGATATTTATGCTAGG	CTTCCAGTATCAACTTGAGTACCTC	TGTCATTCCTAGCATAAATATCCA	3234	12
WIAF-1648	A/G	A	G	ATACACACATGTGCACACAC[A/G]AGAGGCAAGTACAAAAATGT	CACATACACACATGTGCACACA	GGGTTACATTTTTGTACTTGCC	3238	20
WIAF-1652	T/C	C	T	GCAGTAAAGGAACTAATACA[T/C]GTACAGCACCTCAGCACAAA	GGGTCAGCAGTAAAGGAACTAA	TTTGTGCTGAGGTGCTGTAC	3242	17
WIAF-1655	C/T	C	T	TCATACCCCTCAGCTTCCCA[C/T]TGACAGAGCCAGTGTCCTCT	ACCCCTCAGCTTCCCA	GAGGACACTGGCTCTGTCA	3245	

续表

WIAF 命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi 编号	NwLib chrom
WIAF-1660	A/G	A	G	AGATCCCACTTTATTTTTTA[A/G]CTCCAATAAATGTAATTATC	CTCCTTAAGATCCCACTTTATTTTT	GCAGCTGATAATTACATTTATTGGA	1160	1
WIAF-1665	G/A	A	G	AAAAGTAACTATGACTAGAA[G/A]AGGCAGAGGAGTGGGTTGAG	GCACTTAAGGAAAAGTAACTATGACTAGAA	CAACCCACTCCTCTGCCT	1162	2
WIAF-1667	C/G	C	G	CCAGTGGGATGTGGCTTCCC[C/G]AGGTTGCAACCCCAAGGAAG	TGGGATGTGGCTTCCC	CCTTGGGGTTGCAACCT	3249	
WIAF-1670	G/A	A	G	CTGAAATATCTGCTAGTGGG[G/A]AATTTACAACCCACTGACCA	CCCTGAAATATCTGCTAGTGGG	GATGGTCAGTGGGTTGTAAATT	3252	6
WIAF-1671	G/A	A	G	ACAGAGACAAGACATCAATC[G/A]TCTGTTAGCAGCGAGAGAGA	GGTCACAGAGACAAGACATCAA	GTCTCTCTCGCTGCTAACAGA	3253	19
WIAF-1672	G/C	C	G	AGGCCCAGCCCCAGAACCTG[G/C]AGTGCTTCTTTGACGGGGCC	CCAGCCCCAGAACCTG	CCCCGTCAAAGAAGCACT	19250	22
WIAF-1673	C/T	C	T	AAGTGGAGCCCAGAGGTTTG[C/T]TGGGACTCCCAGCCAGGGGA	TGGAGCCCAGAGGTTTG	CCTGGCTGGGAGTCCC	19251	22
WIAF-1674	T/C	C	T	CAGGATTAGTCTTCTCTGTT[T/C]CTTGGTGCAAGTTTGAACCA	GGTTGCATGACAGGATTAGTC	CACTGGTTCAAACTTGCACC	3254	6
WIAF-1676	G/A	A	G	GCAGATCAACAGCTGCTACC[G/A]TTCACCCCAGAGGTGGCTGC	GCAGATCAACAGCTGCTACC	GCCACCTCTGGGGTGAA	1165	10
WIAF-1677	T/G	G	T	TCATGGTCCTTTGTGATCTG[T/G]ACCTCACCCATGTCTCCCAC	ATCATGGTCCTTTGTGATCTG	GGGAGACATGGGTGAGG	3255	1
WIAF-1680	T/C	C	T	CTCATAGAAATCACGGATAG[T/C]ATCACCAGTCTACAGCCACT	CTATGGTTCTTGTAAGGATGTTTCT	AAATTCTGTGATAGATAGTGGCTGT	1168	7
WIAF-1682	G/A	A	G	ACACACTGGCAGGGATTTGC[G/A]TATCTAAACATAGAAAAGGT	CACTGGCAGGGATTTGC	TGGCCCCATAATACCGTAT	1170	21
WIAF-1689	T/C	C	T	ATAAACTCTCTTATGACAAG[T/C]AGTGATTGAACTCTAGTAAA	CCTTTCAGTATAAACTCTCTTATGACAAG	CAAAACCTTTACTAGAGTTCAATCACT	1172	19
WIAF-1690	A/C	A	C	TCTTCCTTTCCCCTTGATCA[A/C]AGTGAAGATATGATAGGGAA	CTTCCTTTCCCCTTGATCA	CTGAATTCCCTATCATATCTTCACT	1173	6
WIAF-1693	G/C	C	G	TTTCCTCCCTCCCCATGACC[G/C]GGCTTCCCGCGGGCACCTGT	CCTCCCTCCCCATGACC	ACGCACAGGTGCCCG	1174	8
WIAF-1694	C/G	C	G	CCTTACAAAAATATTTCTGT[C/G]TAGAGAGGGAAAGAGCTGGT	CCCAGAACCTTACAAAAATATTTC	CACCAGCTCTTTCCCTCTCT	3258	11
WIAF-1696	C/G	C	G	ATTTGCCATCTGTCTGACTC[C/G]CGTCTTCCCGGGGCGTGGGG	CAATTTGCCATCTGTCTGACTC	ACGCCCCGGGAAGAC	3260	6
WIAF-1697	A/C	A	C	AGACCCTTTCCAAGGGAATA[A/C]TACTACACTAAGCCTACACT	AGACCCTTTCCAAGGGAATA	CACAGTACAGTGTAGGCTTAGTGTAGT	3261	16
WIAF-1698	G/A	A	G	AAGATAAAGAGGCAGTGTAA[G/A]AGTAGTATTCTCTACATACC	CATTTTAAGATAAAGAGGCAGTGTAA	GCATCATTGTATACTGTGGTATGTAGA	3262	12
WIAF-1700	G/A	A	G	CAGGTATGGACTATGTACAT[G/A]ACAATACAAGAGGCGTTTGT	GCACTCAGGTATGGACTATGTACA	AAACAAACGCCTCTTGTATTG	2052	2
WIAF-1702	A/G	A	G	TCAACATTACAAAACCCCCA[A/G]TCTTCAAGGAAAGGAGCACA	TCAACATTACAAAACCCCCA	TGTGCTCCTTTCCTTGAAGA	3263	3
WIAF-1703	C/T	C	T	GGAGCACATTACCATGGAGC[C/T]ACAGGACTCCAAAGGACCTC	GGAGCACATTACCATGGAGC	GAGGTCCTTTGGAGTCCTGT	3264	3
WIAF-1706	G/T	G	T	CCAGGCCACTTTGGAGAAAG[G/T]AAGAGAATGCTATTAATCAA	CAGGCCACTTTGGAGAAAG	GGCTTATTGATTAATAGCATTCTCTT	3266	22
WIAF-1709	A/G	A	G	TTTGTAACAGAAACGCCTTA[A/G]GACACTATGTTTGGGAATAT	TGTTTGTAACAGAAACGCCTT	TTCCTATATTCCCAAACATAGTGTC	2053	8
WIAF-1714	C/T	C	T	AATGTTGTCTTTTTTTTTCC[C/T]GGCAAACTTTCTGCTTTCGG	TCGATGGTAATAATGTTGTCTTTTT	CGAAAGCAGAAAGTTTGCC	1182	3
WIAF-1717	G/C	C	G	TGAAAGCCCTCGGCTCGGTC[G/C]TTAGCCCATCTTCCTACATT	GCCCTCGGCTCGGTC	GGTAATGTAGGAAGATGGGCT	2054	16
WIAF-1718	C/A	A	C	ACCTGTGTGTCAGGATGCAC[C/A]TGAAAGCCCTCGGCTCGGTC	CCTGTGTGTCAGGATGCAC	GAGCCGAGGGCTTTCA	2055	16
WIAF-1721	T/C	C	T	TCCATCAAGGACTTTGTTTT[T/C]GTCTCTTCACTCTGCTATTT	TTTCCATCAAGGACTTTGTTTT	TTGTATTATAAATAGCAGAGTGAAGAGAC	3269	5
WIAF-1741	C/T	C	T	AACCTCCTCATTCCTATAAA[C/T]CTTTAACAAAAACAGTTAGC	CCAAACCTCCTCATTCCTATAAA	TGTAAACAGCTAACTGTTTTTGTTAAA	3288	17
WIAF-1742	T/C	C	T	GAAAGGAAAAGCATAAACTT[T/C]AGGATTTCATTGTCTCTTGG	CCAGAAAGGAAAAGCATAAACTT	CCCAAGAGACAATGAAATCCT	3289	10
WIAF-1743	T/C	C	T	GCGTGGAAACCCAATTGTCA[T/C]GTGTATGAACTACAAAAGGA	CGTGGAAACCCAATTGTCA	CCGATCCTTTTGTAGTTCATACA	3290	X
WIAF-1745	T/C	C	T	GATGATCACTGTGCTGCTTG[T/C]GGCTCATGGCAGAGCATTCA	GATGATCACTGTGCTGCTTG	TGCTCTGCCATGAGCC	3292	8
WIAF-1748	A/G	A	G	TGAGATGTCTTTTACCTGAG[A/G]CTAATAAGGATTGAACCAAG	TCCTGAGATGTCTTTTACCTGAG	CTGCTTGGTTCAATCCTTATTAG	3295	1
WIAF-1749	G/A	A	G	AAGGGATTGCTTTATTTAAC[G/A]TGAAAAGCGTGATAGAGGAA	TTTGAAGGGATTGCTTTATTTAA	GTTCCTCTATCACGCTTTTCA	3296	3
WIAF-1753	G/A	A	G	AAGGGAGTGGGGGAGTAAAA[G/A]TGGAAGCAGGGTGACGCATG	GGGAGTGGGGGAGTAAAA	CGTCACCCTGCTTCCA	3300	4
WIAF-1762	T/G	G	T	AATGAAATAGTCTGGCCATT[T/G]GACTAACCAGTTCTACAAAT	AATTAATGAAATAGTCTGGCCATT	TGTGAAATTTGTAGAACTGGTTAGTC	3309	2
WIAF-1763	G/A	A	G	TTAAGTACAAAATGCTAAAC[G/A]GGAGCCGAGCTCTTCCGCAT	TTATTTCCTTAAGTACAAAATGCTAAAC	CGGAAGAGCTCGGCTCC	3310	2
WIAF-1764	A/C	A	C	TACAACACTCCCTTCAGATC[A/C]CAAAAGCTTAACAAATGGTA	AAATATATACAACACTCCCTTCAGATC	CGTTTTACCATTTGTTAAGCTTTTG	3311	9
WIAF-1765	G/A	A	G	AATGTTAATCGTAACATACT[G/A]GAAAGCTGTTACAGTAGAAG	CCCCAAATGTTAATCGTAACA	GCTAACACTTCTACTGTAACAGCTTTC	3312	3
WIAF-1767	T/G	G	T	TATAATGGGGATTTTCTGCT[T/G]AACTGCCCACTGATTCTTAC	CATTTTTATAATGGGGATTTTCTG	TGTAAGAATCAGTGGGCAGTT	3314	11
WIAF-1768	A/G	A	G	CTGTAGACATCTAACATTAG[A/G]TAGCCTTCAGAATTGCAAGT	AAAGATTCCTGTAGACATCTAACATTAG	CACTTGCAATTCTGAAGGCT	3315	1
WIAF-1772	A/G	A	G	GTTTGATTTGGCTTCCCTAT[A/G]GATTCAGGACCCATAACTCT	TGTTTGATTTGGCTTCCCTA	CAAGAGTTATGGGTCCTGAATC	3319	7
WIAF-1774	G/A	A	G	AATAAGCAGCTAATGGCAAT[G/A]CTAGTGGTCTTCCCAATTCA	CAAATAAGCAGCTAATGGCAA	TGTGAATTGGGAAGACCACT	3321	2
WIAF-1775	T/C	C	T	TCCTAGTCACCAAGCATACT[T/C]TCCTGGCTCCCCAAGTACTT	TTCTCCTAGTCACCAAGCATACT	ACTTGGGGAGCCAGGA	3322	18
WIAF-1778	A/G	A	G	CACCTGACTCGGTGCTTTAC[A/G]TACATTACCTCACAGCCAGG	CACCTGACTCGGTGCTTTAC	CCCTGGCTGTGAGGTAATGT	3325	14
WIAF-1779	C/T	C	T	GGGTGGGGTACTGATTTATC[C/T]TAGATCCAAATAAAGCATGC	GGGGTGGGGTACTGATTTATC	CTGCATGCTTTATTTGGATCT	3326	1
WIAF-1780	A/C	A	C	TACAGCCAACATCACTGTTT[A/C]ATTCCAGAACATTTTCAACA	TGTACAGCCAACATCACTGTTT	GAGATGTTGAAAATGTTCTGGAA	3327	6
WIAF-1781	C/G	C	G	CTTCCTGTCTGATTTTCCAA[C/G]ACTATGTTTAATGTATGACT	TCATCTGAGAATAAACTTCCTGTCT	CATTATAGGTACTGAGTCATACATTAAACA	3328	7
WIAF-1785	C/T	C	T	GACCTAGCTGGCTTTGTAGT[C/T]GTTCAGGCCCATTGAAATAG	CAGACCTAGCTGGCTTTGTAGT	CTATTTCAATGGGCCTGAAC	3332	7
WIAF-1786	A/G	A	G	CCAACCTACTTTGGAGCCCT[A/G]AGGAGTTTTTAGAGAAAGCT	CCAACCTACTTTGGAGCCCT	TCCAGCTTTCTCTAAAAACTCCT	3333	3
WIAF-1789	A/G	A	G	AATACACTGACAGGTAAATA[A/G]TATAACATTAGAAAAGCAAA	TGCAGTAATACACTGACAGGTAAAT	GTTAAAAGAATTTTGCTTTTCTAATGTT	3336	3
WIAF-1793	G/A	A	G	AGTTTCACAATATAGGTAGC[G/A]ATAACCAGGTCTCACTTTCC	CCCAGAGTTTCACAATATAGGTAGC	GGGAAAGTGAGACCTGGTTAT	3340	3

续表

WIAF 命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi 编号	NwLib chrom
WIAF-1796	A/T	A	T	CCACACATTTCAGTCCAAGA[A/T]AACCTTCCTCAAATTTTCCA	CCACACATTTCAGTCCAAGA	GATGGAAAATTTGAGGAAGGTT	3343	5
WIAF-1799	T/C	C	T	GGACATTTGGATGGTGACTT[T/C]CCTGGGTGGTTCCCCATAGA	GGACATTTGGATGGTGACTT	GGGGAACCACCCAGG	3346	10
WIAF-1804	A/G	A	G	ATCACATCTTCAGGATAGGT[A/G]ATAACAGTGTGAAGGGTGTG	AGTGTATCACATCTTCAGGATAGGT	GCACACCCTTCACACTGTTA	3351	5
WIAF-1805	T/C	C	T	ACTCTGGTTTTATAAAGCTA[T/C]AGGACAGAGCAGAGATGGAA	CCACACACTCTGGTTTTATAAAGC	TTCCATCTCTGCTCTGTCCT	3352	12
WIAF-1810	G/A	A	G	GAGTAAAGATGGCGCTAGAA[G/A]GTATCTGTTATAGAAACGAT	AAGAGTAAAGATGGCGCTAGAA	CAAAATGAAGTATCGTTTCTATAACAGA	3357	7
WIAF-1811	G/T	G	T	CCAAGGGATTAACGTATAGG[G/T]TCTTAAACAAGGGGATCCCC	GCCAAGGGATTAACGTATAGG	GGGGATCCCCTTGTTTAAGA	3358	18
WIAF-1812	G/A	A	G	CAGCTTTCGTTAGGCTAGTT[G/A]GCTGAGCCATTGTATGCGGA	AAACAGCTTTCGTTAGGCTAGTT	CGCATACAATGGCTCAGC	3359	16
WIAF-1816	A/G	A	G	CGTCCCGTCCCGCCAGCCCT[A/G]TCGGCCTCGTCACTGGCCTT	GTCCCGTCCCGCCAG	CCAGTGACGAGGCCGA	3363	22
WIAF-1819	A/G	A	G	GACCTGCAGAACTTAAACAC[A/G]GAGCATTTATTGTTAGAAAG	GGAGACCTGCAGAACTTAAACA	GCCCTTTCTAACAATAAATGCTC	3366	3
WIAF-1827	A/G	A	G	TGTAAAATAACAATAACGTT[A/G]AAGGCAAAAGCAAGATTCTG	CATTTCTTTGTAAAATAACAATAACGTT	CAGAATCTTGCTTTTGCCTT	3374	5
WIAF-1834	G/A	A	G	ATTCATAAACATATCAACCA[G/A]TAGCATTAACCCATTTTATT	GAGTGATTCATAAACATATCAACCA	AGGAAATAAAATGGGTTAATGCT	3381	3
WIAF-1842	G/A	A	G	TAGTCACTTTGAAATGTAAC[G/A]AATGGTACTACAACCAATTC	CAGCATTAGTCACTTTGAAATGTAA	TTGGAATTGGTTGTAGTACCATT	3388	1
WIAF-1846	A/G	A	G	TGCACAGCATGGCTAAAACG[A/G]TAAAGATGGGAATCAACAAA	GCACAGCATGGCTAAAACG	GGTATTTGTTGATTCCCATCTTT	3391	8
WIAF-1847	C/T	C	T	ATTATTTCAGGCCAATGTGT[C/T]TGTTGGGTCTGAGATTTGAT	GAGATTATTTCAGGCCAATGTG	AATCAAATCTCAGACCCAACA	3392	8
WIAF-1849	C/T	C	T	CCCAACAACATCCTCTGCCA[C/T]ACACAACAAAACGTAAGTTT	CCAACAACATCCTCTGCCA	CCAAATGAAACTTACGTTTTGTTG	3394	11
WIAF-1851	G/A	A	G	AACTTTGGGAAAAGGTGTAA[G/A]ACAGTAGCCCCATCACATTT	AAATAAACTTTGGGAAAAGGTGT	CAAATGTGATGGGGCTACTG	3396	X
WIAF-1853	C/T	C	T	AGCTGTTCAGGACTGGACTC[C/T]GGTCCCTTTATTGAGACTGA	AAGCTGTTCAGGACTGGACTC	TGTCAGTCTCAATAAAGGGACC	3398	20
WIAF-1854	G/A	A	G	CATTAAGCATCACTGTCACT[G/A]GCTAACTCCTCAAATCAACA	GGAACAAACACATTAAGCATCA	GGTATTGTTGATTTGAGGAGTTAGC	3399	1
WIAF-1856	C/T	C	T	TTTTTACAGAGGTAGCACAA[C/T]TGATTCCAACACAAAACCCC	TTCATTTTTACAGAGGTAGCACAA	GGGGTTTTGTGTTGGAATCA	3401	21
WIAF-1857	T/C	C	T	TTTCAACAGGAAAAGCCATG[T/C]ATGACATTCAAAACACCGTA	TTCAACAGGAAAAGCCATG	AATAATACGGTGTTTTGAATGTCA	3402	21
WIAF-1858	C/G	C	G	CTGAGTCACGCTGAGGAGAG[C/G]CTTCACTCAGGAGTTCATGC	TGAGTCACGCTGAGGAGAG	AGCATGAACTCCTGAGTGAAG	3403	13
WIAF-1861	C/T	C	T	ATCGGGTTGTCCAGCCAACA[C/T]GGAGGTGATTTTGGTGGGGA	GGGTTGTCCAGCCAACA	CCCACCAAAATCACCTCC	3406	15
WIAF-1862	A/G	A	G	TCCTCAGAACTTCTCAGCCT[A/G]GTAGCACAAGTGGATGCTTG	ATCCTCAGAACTTCTCAGCCT	TCAAGCATCCACTTGTGCTA	3407	11
WIAF-1866	C/T	C	T	AGTAAAAAGAAACTCATGAC[C/T]TTCTCCTTGGACTCGCTCCT	CACGAGTAAAAAGAAACTCATGAC	GGAGCGAGTCCAAGGAGAA	3411	11
WIAF-1869	C/T	C	T	TACAACAGAGGAGATAATCT[C/T]CAGGATGCCTGTGAAGAAAG	AGGAGATACAACAGAGGAGATAATCT	TTTCTTCACAGGCATCCTG	3414	5
WIAF-1870	T/A	A	T	TGATGAGACTAGAGTGACAG[T/A]GTTTCAGAACCCAAATGTCC	TTTTCTGATGAGACTAGAGTGACAG	AGGACATTTGGGTTCTGAAA	3415	16
WIAF-1871	G/A	A	G	CTGCTAAGCTGTATTTCAGA[G/A]GAATGTCACAATCATACCAC	GTAGCTGCTAAGCTGTATTTCAGA	CCAGTGGTATGATTGTGACATTC	3416	10
WIAF-1872	G/A	A	G	TGCACAGAAAGCATACTTCT[G/A]TGGCTTTGTTACACGGGTTT	TGATATGCACAGAAAGCATACTTC	ACCCGTGTAACAAAGCCA	3417	X
WIAF-1883	A/G	A	G	TGCTCTCGGAGGATGCCTAG[A/G]AGATGTTGGGAACAGAAGAA	GCTCTCGGAGGATGCCTAG	TCCCCCTTAAACTCAGTTTATTT	3428	1
WIAF-1885	T/C	C	T	ATTACTAGCCTAGATCCTAA[T/C]TGAGGACATGAGATTTATTG	AAACATCATTACTAGCCTAGATCCTAA	CCTTCAATAAATCTCATGTCCTCA	3430	8
WIAF-1886	T/G	G	T	AAAGTCTGGGATGACTTTCC[T/G]ATTCTACATCAAGTAGAACC	CCTAAAGTCTGGGATGACTTTCC	TTGGCTTAGGTTCTACTTGATGT	3431	2
WIAF-1895	C/T	C	T	GATACAGGCCATATTTCCCA[C/T]ATAGGACTCTAGTTCTAGAA	GATACAGGCCATATTTCCCA	CAAGGCTTTCTAGAACTAGAGTCC	3440	14
WIAF-1897	C/A	A	C	GTTGGCAGACAACACACTAG[C/A]ATTTTCACGGGTGTGGGCAC	GGTTGGCAGACAACACACTAG	CCCACACCCGTGAAAAT	3442	8
WIAF-1898	C/T	C	T	TGTATTTACACGAGGGTAGA[C/T]GGCAGATGCCTGACAGAGAG	ACGTGTATTTACACGAGGGTAGA	CTCTGTCAGGCATCTGCC	3443	8
WIAF-1899	A/C	A	C	AAGAATTTTGCATGGCGATT[A/C]AAATAGAAAACCTATAAATG	GAATTTTGCATGGCGATT	AGACCTGCTTTTTCTACATTTATAGG	3444	9
WIAF-1900	G/A	A	G	AGATAAACCAAAATGATTGA[G/A]TATGATAAAGAATTTTGCAT	GGAAGATAAACCAAAATGATTGA	GCCATGCAAAATTCTTTATCA	3445	9
WIAF-1904	C/T	C	T	AGTCTGGTCACTTTGGGGCC[C/T]GGCGTGGGCAGAGCCCACTG	TGGTCACTTTGGGGCC	GGCTCTGCCCACGCC	3449	22
WIAF-1906	A/G	A	G	TTTCTGGTCACTGCAGCCCC[A/G]TCTGTATTAGGGAGCACCCC	TGGTCACTGCAGCCCC	GGGGTGCTCCCTAATACAGA	3451	3
WIAF-1909	A/C	A	C	AACCACTGTCCCTGTAATGT[A/C]CTGCTGAGAGTCCACATTTT	CAACCACTGTCCCTGTAATGT	CAAAATGTGGACTCTCAGCAG	3454	16
WIAF-1910	T/C	C	T	CTTTTGTTTCTACCCCCTAT[T/C]CATTACAGTCAAATTAACAG	ATCCTTTTGTTTCTACCCCC	TTGCCTGTTAATTTGACTGTAATG	3455	5
WIAF-1918	T/C	C	T	AACCAGATTTTCTCAGGCCT[T/C]TTTTGGATACCTTTAGTAGT	AACCAGATTTTCTCAGGCCT	ACAAGAGGGTTTGACAAAAGAG	3463	12
WIAF-1919	G/A	A	G	CTTAAACTCAGTTTCCGCTG[G/A]TGCTATGTAAAGCATCCACG	GACTTAAACTCAGTTTCCGCTG	CCATCGTGGATGCTTTACAT	3464	9
WIAF-1923	A/G	A	G	GAAAATGCCACGTCTCTGAC[A/G]GCGATTTACCTGACATGTGT	GAAAATGCCACGTCTCTGAC	GACACATGTCAGGTAAATCGC	3468	9
WIAF-1925	G/T	G	T	AAACTACATTTGGGTTTTGG[G/T]AAGTCCCCTGTAATGATGAA	AGATAAACTACATTTGGGTTTTGG	GATTCATCATTACAGGGGACTT	3470	4
WIAF-1928	T/C	C	T	ATACCAAGTGTACATACTGT[T/C]CACATGATTTATGGCTGTTG	GCTATTAGTTACATCATACCAAGTGTACA	TCAACAGCCATAAATCATGTG	3473	3
WIAF-1929	A/C	A	C	AAGGGAAAGCAGACCTGGGG[A/C]CCACGGGCAATCACATGAGA	GGAAAGCAGACCTGGGG	ATGTGATTGCCCGTGG	3474	2
WIAF-1930	C/G	C	G	TCATGGAAATAGGCCTGGAG[C/G]ACAGGATTTGGCTGAGGCTT	CATGGAAATAGGCCTGGAG	GCCTCAGCCAAATCCTGT	3475	2
WIAF-1934	C/T	C	T	CACAGGAATCAGCAGCCTGA[C/T]TGTTGCACTTGTCCAAACAC	CAGGAATCAGCAGCCTGA	TGTTTGGACAAGTGCAACA	3479	1
WIAF-1938	G/A	A	G	TGTCTCACCTGGACTTAAGC[G/A]TCTGGCTCTAATTCACAGTG	CCTGTCTCACCTGGACTTAAGC	GCACTGTGAATTAGAGCCAGA	3483	3
WIAF-1940	A/G	A	G	CAAAAAGGACTCTGCATTGT[A/G]ATTAAGTTTATTAATCAGCT	GCAAAAAGGACTCTGCATTG	GCTAAAGTCAGCTGATTAATAAACTTAA	3485	19
WIAF-1942	G/A	A	G	AAGCAGCTCATTTTGATTAC[G/A]GGTATACATGAAGTAAAATT	CAATAAGCAGCTCATTTTGATTAC	ACTTCATGAATTTTACTTCATGTATACC	3487	2
WIAF-1943	T/C	C	T	TGTCCTAGCTAATGAATGCA[T/C]AGAGTATTGCCTGCAAAATA	GTTGTCCTAGCTAATGAATGCA	AATTATTATTTTGCAGGCAATACTC	3488	21
WIAF-1945	G/A	A	G	GTAAGAAATCAGGGATAAGA[G/A]CTGAGGAACAAGAGGGATAT	TCCTGTAAGAAATCAGGGATAAGA	AGTCCTGACTCACTGCCTACA	3490	15
WIAF-1946	T/C	C	T	TTAGCAGCAGCAGAATGCAC[T/C]TTTGCAGAAACACACAAACA	GCAGCAGCAGAATGCAC	GTGTTTGTGTGTTTCTGCAAA	3491	3
WIAF-1947	C/T	C	T	TGCTCCCCACAATTTTGATT[C/T]GGTGGCTTCATAAGGGACCC	GCTCCCCACAATTTTGATT	GGTCCCTTATGAAGCCACC	3492	5
WIAF-1948	C/T	C	T	GTTGGGGGAGGTAGGAGACT[C/T]GGACCGGCAGCCCTGGCTCC	TTGGGGGAGGTAGGAGACT	CAGGGCTGCCGGTCC	3493	19

WIAF 命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi 编号	NwLib chrom
WIAF-1949	A/G	A	G	AGACCCTTAGCTGATCTCAT[A/G]AAGTCCACCTCATGAAGGAG	TGAGACCCTTAGCTGATCTCA	TCTCCTTCATGAGGTGGACTT	3494	11
WIAF-1950	C/G	C	G	ACACCATGTCCCTAAAATGT[C/G]ATTCAACATATATGCACACC	AAACACCATGTCCCTAAAATGT	CGAAGGTGTGCATATATGTTGAA	3495	11
WIAF-1955	G/A	A	G	TCAGAAGCTGATAAACGTGG[G/A]CTTACACCTTTAGCACGGAT	GTCAGAAGCTGATAAACGTGG	AAACTATCCGTGCTAAAGGTGT	3500	6
WIAF-1956	C/T	C	T	TTTCTGCCAGCTTACAGGCT[C/T]ACAGAAGAATGAGACACTTA	TCTGCCAGCTTACAGGCT	GCGTAAGTGTCTCATTCTTCTGT	3501	1
WIAF-1958	A/G	A	G	AGGGAAACTTATAACCTCAC[A/G]CGCTTGTTTCACAAAACAAC	GTTGAGGGAAACTTATAACCTCAC	TGTTGTTTTGTGAAACAAGCG	3503	4
WIAF-1959	T/C	C	T	AACCACCTTCAGAAATTCTA[T/C]AAAACACTAGCAACTTCCTT	GATTTAACCACCTTCAGAAATTCT	TGATAAAAGGAAGTTGCTAGTGTTT	3504	6
WIAF-1960	C/T	C	T	ATAGGGACTTCAGAGTAATA[C/T]GGTTTATGTCAGTTTTCCAG	CTACTGACATAGGGACTTCAGAGTAA	TCCTGGAAAACTGACATAAACC	3505	7
WIAF-1964	T/C	C	T	TCAGAGCCAGGGCTAGAGGA[T/C]GCACGGTGGCTAGAGCCCGC	AGAGCCAGGGCTAGAGGA	GGCTCTAGCCACCGTGC	3509	2
WIAF-1969	C/A	A	C	CCTGGGCACATTATCCCCCT[C/A]GGGTAAACCAGGACTATTGC	GGGCACATTATCCCCCT	TGCAATAGTCCTGGTTTACCC	3514	20
WIAF-1970	G/A	A	G	GGAAGGAAGGGGCGGTCATT[G/A]GGTGATGGCTTCTGGCTCTC	GGAAGGGGCGGTCATT	AGCCAGAAGCCATCACC	3515	1
WIAF-1971	A/G	A	G	AGGAGCCATTGTTGGGGTTA[A/G]ACTGTCCTGAACAAAATCTA	GGAGCCATTGTTGGGGTTA	GCCTAGATTTTGTTCAGGACAG	3516	16
WIAF-1974	A/G	A	G	CAGAAATGCAGGAGGGTGGC[A/G]AGAGGGGCCGAGATTGGGTG	AATGCAGGAGGGTGGC	CAATCTCGGCCCCTCT	3519	22
WIAF-1977	A/G	A	G	CAGCTTCTACATTCTGAATA[A/G]AGTACATAATGGGATTTAAG	CAAGTCAGCTTCTACATTCTGAATA	TAAAGATTTACTTAAATCCCATTATGTACT	3522	7
WIAF-198	C/T	C	T	GAGGCCTAAAGGAATGGGAA[C/T]GTGTTGGTGGTCGCTTGATA	GGCCTAAAGGAATGGGAA	TCAAGCGACCACCAACAC	19249	1
WIAF-1980	T/C	C	T	TGCAGGTGCACCGAAAGGAC[T/C]TGGGAGATAAAATTCAAAAA	CAGGTGCACCGAAAGGAC	GAGCAGCACATCACACTTTTT	3525	12
WIAF-1988	A/G	A	G	AACCATAGGAGCCAAAAGTC[A/G]GACAAACAGAAGAAGGCACA	GAACCATAGGAGCCAAAAGTC	GTGTGCCTTCTTCTGTTTGTC	3533	12
WIAF-1991	A/G	A	G	AGCCATCTTTCCGGAAGCTC[A/G]TGGAGCACAAGCAGAACTCG	CCATCTTTCCGGAAGCTC	GAGTTCTGCTTGTGCTCCA	3536	7
WIAF-1992	G/A	A	G	TTGTGATGAAGAACATGATG[G/A]TCACTAGTAGGTAACTTTCT	ACATTGTGATGAAGAACATGATG	AATGACACAGAAAGTTACCTACTAGTGA	3537	8
WIAF-1996	A/G	A	G	CTTCTCTTTAGGTAATTTGC[A/G]TAAGAACAATAAAAGCATTT	GGATTAATCTTCTCTTTAGGTAATTTGC	GACTTTTAAAATGCTTTTATTGTTCTT	3541	17
WIAF-1997	G/C	C	G	AAACTTTAGATATCTCCCAT[G/C]TTCCACAGAATCAAATATAT	TGTTCTCTAAACTTTAGATATCTCCCA	CAAGAAATATATATTTGATTCTGTGGAA	3542	14
WIAF-1999	A/G	A	G	AATGGATCCTTATCTGCACA[A/G]CCATTGAAGAAAAAAAAAAA	AAATGGATCCTTATCTGCACA	AGTTTCAGTTTGCATGAATTTTT	3544	1
WIAF-200	A/G	A	G	CCTAGTGGCATTAAGGATGC[A/G]GTAGGATGTCCACTTTTAGT	CCTAGTGGCATTAAGGATGC	TTGCTACTAAAAGTGGACATCCT	2570	1
WIAF-2006	A/G	A	G	TCAAGCACAGCCCTGCCCCC[A/G]TCTTGAGATTCAGAATCCAG	CACAGCCCTGCCCCC	CCCTCTGGATTCTGAATCTCAA	3551	20
WIAF-2007	T/C	C	T	TGTCACTCTAGCATCTGGAA[T/C]GCTCCGTTGTATATTCAGGA	GCTGTCACTCTAGCATCTGGAA	CCTCCTGAATATACAACGGAGC	3552	1
WIAF-201	C/T	C	T	TGAGAGCTTAAAACAACACT[C/T]ATTTGTTATTTCACAGTTCA	AGACTTGAGAGCTTAAAACAACACT	GCCTACTGGACCTCTAAACTACTGA	2571	3
WIAF-2012	C/T	C	T	AAAACCACAGGGATTCCGGA[C/T]GCCAGACCCCATTTTATACT	CCACAGGGATTCCGGA	AAGTATAAAATGGGGTCTGGC	3557	15
WIAF-2013	T/A	A	T	ATCAATAGGAGAGGATTGGC[T/A]TTTGAATTCAGAGCAAAGCC	AATCAATAGGAGAGGATTGGC	GGCTTTGCTCTGAATTCAAA	3558	10
WIAF-2014	T/C	C	T	AGATCTGTCTGGCATTCTTT[T/C]GTGGGGCTGTTTTTCCAAGG	GAAGATCTGTCTGGCATTCTTT	TGGAAAAACAGCCCCAC	3559	1
WIAF-2016	G/A	A	G	GAGTCCTGCCCAGCAGCCTC[G/A]GTGGCCAAGCCCAGACACTC	CCTGCCCAGCAGCCT	TCTGGGCTTGGCCAC	3561	10
WIAF-2019	G/A	A	G	AGAGCAGAGTTCCTGCCCTC[G/A]GTGTGCGGGGGGAGAGAGGG	AGCAGAGTTCCTGCCCTC	CCTCCCTCTCTCCCCC	3564	20
WIAF-2021	T/C	C	T	AATATAGAGAATCCTGCACT[T/C]CCCAAGTCTCGTCGCACAGG	CACAAAAATATAGAGAATCCTGCA	TGCGACGAGACTTGGG	3566	6
WIAF-2023	T/C	C	T	TGTAGAATTTGTGTCGATGC[T/C]CTGTGTCCTCCGTCCTTCCC	GCTGTAGAATTTGTGTCGATGC	GGAAGGACGGAGGACACAG	3568	5
WIAF-2029	C/A	A	C	ACAAAGGCAACTATGTGCAG[C/A]AACAATCTGATGGGCAGTCC	ATACAAAGGCAACTATGTGCAG	CAAGAAGTTTGGACTGCCC	3574	15
WIAF-2041	C/T	C	T	AAAAATCCAAGCTTACAACT[C/T]GTCCTTTACCTGATACATTT	GCTAGACCTAAAAATCCAAGCTT	TGGAATAAATGTATCAGGTAAAGGA	3583	14
WIAF-2043	C/T	C	T	ATCTTATGTCTGGCTCTATT[C/T]ATCCCTATTCTCTGATCTTA	GATCTTATGTCTGGCTCT	ATAAGATCAGAGAATAGGGA	784	11
WIAF-2046	A/C/G	A	C	GTGATTTTCCACACATTGTA[A/C]AGTGAAAGCTCTTCAGCTTG	TGGTGATTTTCCACACATTG	TCCAAGCTGAAGAGCTTTCA	3585	11
WIAF-205	G/A	A	G	AGTGGAACCCTGAGACTTTA[G/A]ATCTGCAAAGGGGTTTAATA	CTAATAGTGGAACCCTGAGACTTT	CATTATTAAACCCCTTTGCAGA	2575	5
WIAF-2057	C/G	C	G	AACACACAAATACAGAGATT[C/G]AATTCAGGAGCCAGTTTCTA	TCCCCCACAACACACAAATA	CCTAGAAACTGGCTCCTGAA	1081	22
WIAF-2059	C/T	C	T	TGGCACCGTTTGGCTCATTC[C/T]ACTCTCCCTGGGTCTTATTG	GCACCGTTTGGCTCATTC	GTCAATAAGACCCAGGGAGAG	3590	3
WIAF-2061	C/T	C	T	GCTCCACCTTTAATCTGTTT[T/C]ATACCTTGCTTATTAAATGA	GGCTCCACCTTTAATCTGTTT	TCGCTCATTTAATAAGCAAGG	3591	19
WIAF-2062	A/G	A	G	CCTTCTAACTGAGGAACTAC[A/G]TTATACTGGAATCATGTGAA	CCAACCTTCTAACTGAGGAACTAC	TTTAGAATGTCTTCACATGATTCC	789	13
WIAF-2063	C/G	C	G	GTCTCCTGAGTGTCTGAATG[C/G]GCCAGGTGGCTAAGTGCTGG	GGTCTCCTGAGTGTCTGAATG	CAGCACTTAGCCACCTGG	1083	19
WIAF-2064	C/T	C	T	CAGGTGGCTAAGTGCTGGGG[C/T]TCTGGGGTCAGGCTGCCTGG	CCAGGTGGCTAAGTGCTGG	GGCAGCCTGACCCCA	1084	19
WIAF-2067	A/G	A	G	TTGTGTATTGGCATTTTAAA[A/G]GTACCATTCCATTTTCTTCT	CAATTATTGTGTATTGGCATTTTAAA	GCCAGAAGAAAATGGAATGG	1085	8
WIAF-207	G/A	A	G	AGACTTCAAAAGCAATTCAC[G/A]CTTCCAGAATACAAAGTACT	GACAGACTTCAAAAGCAATTCA	AGGTTTGAAAATATGTATTAAGTACTTTGT	19227	6
WIAF-2072	A/G	A	G	ATGAGCTTGTGCACCATTAG[A/G]TCCTGCTGGGTGTTCTCAGT	AATGAGCTTGTGCACCATTAG	TGAGAACACCCAGCAGGA	1185	1
WIAF-2074	C/T	C	T	CTCCCTCCCTCCAAAACAAA[C/T]GAACAAAAATAAAGAAAGAA	CCCTCCCTCCAAAACAAA	ATGGGTTTTCTTTCTTTATTTTTG	1187	3
WIAF-2077	G/T	G	T	CAAACTGTCACCTTTGTAGC[G/T]CTGGGTCAAAGTCTAAAGAG	CCAAACTGTCACCTTTGTAGC	TCCTCTTTAGACTTTGACCCAG	1190	15
WIAF-2078	G/T	G	T	GGAGTTTGGGTGTGCACTGG[G/T]TGTCTTTCAACTGGGTGGAA	AGTTTGGGTGTGCACTGG	TCCACCCAGTTGAAAGACA	1191	1
WIAF-2079	C/T	C	T	GAGTGCTCAGTGATGTGAAG[C/T]ACACAGGAGTCCCTCAGGGC	CTAAATGAGTGCTCAGTGATGTG	CCCTGAGGGACTCCTGTGT	1192	9
WIAF-2080	G/A	A	G	CAGCCAACAGCAAACAGCCC[G/A]AGCAGGAATCGGCACATAGT	CCAACAGCAAACAGCCC	TGTGCCGATTCCTGCT	1193	2
WIAF-2081	C/T	C	T	GCTAGGCAGTGAACACATCA[C/T]GTATGCAATGAGAAAATAAC	GCTAGGCAGTGAACACATCA	CCTACCAGTTGGTTATTTTCTCA	1194	3

续表

WIAF 命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi 编号	NwLib chrom
WIAF-2084	G/C	C	G	GAATTCAGAAAATTGTGTGT[G/C]TGGGAGGCAGGGTAGCAAGA	CCAAAGAATTCAGAAAATTGTGTG	GCTACCCTGCCTCCCA	1197	8
WIAF-2086	G/A	A	G	GTGTCCCAAGGGAGACGGCC[G/A]GGCTCACACATCCCATCAAA	CCAAGGGAGACGGCC	TGATGGGATGTGTGAGCC	1199	7
WIAF-2087	G/A	A	G	GTGCACATCCACACAGGCAC[G/A]TAACATACACAGTACTGTCT	GCACATCCACACAGGCAC	AGGTGTTGATAACTAGACAGTACTGTG	1200	5
WIAF-209	G/A	A	G	GGTTTTTCACCCTATTCTTC[G/A]TAGACCCTGGGGAGAAAAAA	TGAGGTTTTTCACCCTATTCTTC	TTTTTCTCCCCAGGGTCTA	2577	5
WIAF-2091	T/C	C	T	GGCACAAAAAACCCCAATGA[T/C]CCTATTTCCAAGAATGTATC	CCACCCTGGCACAAAAA	TCTGGATACATTCTTGGAAATAGG	1204	4
WIAF-2092	G/A	A	G	AAGAGAGAGGATATGGTCCC[G/A]TTGCTGACTCCATGTGTTGC	GGAAAGAGAGAGGATATGGTCC	GCAACACATGGAGTCAGCAA	1205	8
WIAF-2094	C/A	A	C	CGGGAGGCTTCACGTCCTCG[C/A]CCGTGGTCCCTGGGTGGCCT	GAGGCTTCACGTCCTCG	ACCCAGGGACCACGG	1207	9
WIAF-2095	G/C	C	G	TGGCTCTTCCTTCTGTGCGT[G/C]TTCGGGAGGCTTCACGTCCT	GGCTCTTCCTTCTGTGCGT	CGTGAAGCCTCCCGAA	1208	9
WIAF-210	C/T	C	T	ACACATATGAGTGAAATTTC[C/T]GGGGCATGGGAAATACATCT	TTTTGTTAGTCTATATTCACACATATGAGT	TGTATTTCCCATGCCCC	2109	18
WIAF-2100	A/G	A	G	CCCTCAGGCTGTCCTACTCA[A/G]TGTGGTTTGCTAGCCTCACT	CCCTCAGGCTGTCCTACTCA	GTGAGGCTAGCAAACCACA	1213	8
WIAF-2101	C/T	C	T	GGCCGAGGAGGAAGGGGCTC[C/T]GTGCACTTGCAGGCCACGTC	CGAGGAGGAAGGGGCTC	TGGCCTGCAAGTGCAC	1214	19
WIAF-2103	A/G	A	G	TGCCCCTGGGACTGGAGCAG[A/G]CTTGGGTGAGCTCTAGGTGG	CCCTGGGACTGGAGCAG	CCACCTAGAGCTCACCCAAG	1216	6
WIAF-2104	G/A	A	G	GACCACGACTACGCTCTGCC[G/A]GTGGGAAAGCAGAAGCAGGA	CCACGACTACGCTCTGCC	CTGCTTCTGCTTTCCCAC	1217	11
WIAF-2105	A/G	A	G	AGAATCATGGGGCAAAAGTC[A/G]CTATGGGGCCAGACTGAGGT	GAATCATGGGGCAAAAGTC	CCTCAGTCTGGCCCCATAG	1218	2
WIAF-2106	C/T	C	T	GGGGGCTGGGCACCTTTTCT[C/T]CAGCCACAGGCCCCTGAGGA	GGCTGGGCACCTTTTCT	AGGGGCCTGTGGCTG	1219	6
WIAF-2108	A/G	A	G	AGGGGCTCCACAGAGAGAGC[A/G]TAAGGGGAAGACTTTTATAG	GGGCTCCACAGAGAGAGC	TGTCCTATAAAAGTCTTCCCCTT	1221	5
WIAF-2112	T/C	C	T	TTTTCTGTGGTGTCAGCAAA[T/C]GCCCCTTTATTTTAAATGAT	TTTTCTGTGGTGTCAGCAAA	GGAATCATTTAAAATAAAGGGGC	1225	20
WIAF-2114	C/T	C	T	CAGAATGGCAGCACCACTGG[C/T]ATGGCGATGGTGCAGGTGGG	ATGGCAGCACCACTGG	CCTGCACCATCGCCAT	1227	4
WIAF-2116	A/G	A	G	ATTTTTTAGACCGTGATTTC[A/G]AAAGAAACAATAAATGTGGA	CGGCTATAGTCTGAGTCATATTTTT	CTTTCTAATCCACATTTATTGTTTCTT	1229	3
WIAF-2119	G/A	A	G	TGATTGAAAAAAGTCAAAAC[G/A]TGAAGAAGCATCAAAGCCAA	CTTGAATATGAAACTGTGATTGAAA	TGGCTTTGATGCTTCTTCA	1232	20
WIAF-2120	A/G	A	G	AGAGGTTGAAGAAGTGCTGA[A/G]AAATATATTTAAGATTTCCT	GAAGAGGTTGAAGAAGTGCTGA	TCTCCCCAAGGAAATCTTAAA	1233	8
WIAF-2122	C/T	C	T	CTACAAATCTGTTGTGTGCT[C/T]GCCGCGTGACTCAGCTAATG	TGCTACAAATCTGTTGTGTGCT	GCTGAGTCACGCGGC	1235	1
WIAF-2128	A/G	A	G	TTTCCTTTGTGTACATTTCT[A/G]TATATTATTTTACTTCTTCT	GCAGATATGACTTTCCTTTGTGT	GGCATTTTCAGAAGAAGTAAAATAA	1241	20
WIAF-2129	C/T	C	T	TGAGCCTGATATTCACACTA[C/T]CTACATTCCCTCCAGTATAA	GATTTGAGCCTGATATTCACACT	TCCTATTATACTGGAGGGAATGTAG	1242	20
WIAF-2130	G/C	C	G	AGTACCTACCTGGGAGGTTG[G/C]TGTACTTGGCTTAAGTACTT	CAGTACCTACCTGGGAGGTTG	GCATGAAGTACTTAAGCCAAGTACA	1243	2
WIAF-2133	A/G	A	G	GCAAGTCATGTGTTAGGTCC[A/G]CTCCCTTGCATGAAATGTGG	TGCAAGTCATGTGTTAGGTCC	CACATTTCATGCAAGGGAG	1246	9
WIAF-2134	T/G	G	T	AAAATAATAAATGAACGTGC[T/G]GATAAACATTCTTCTTATGG	TCCAAAATAATAAATGAACGTGC	TGGAACCATAAGAAGAATGTTTATC	1247	22
WIAF-2135	A/G	A	G	GAGTAAAGGATGCAATCACG[A/G]CTCACTGTAGCCTGGACCTC	GGAGTAAAGGATGCAATCACG	GAGGTCCAGGCTACAGTGAG	1248	2
WIAF-2136	T/C	C	T	CCGCACAGCACTTTGTGTCT[T/C]GCTTTGAGCACTTGCCACTC	CGCACAGCACTTTGTGTCT	GTGGCAAGTGCTCAAAGC	1249	16
WIAF-2137	T/A	A	T	TCTTTGCTGGGTAACACTGA[T/A]CAAGTTGCTTAACCTTTGTG	CTCTTTGCTGGGTAACACTGA	GGTTTCACAAAGGTTAAGCAAC	1250	18
WIAF-2138	T/C	C	T	ATATTACTATTTAGTCTAAG[T/C]TTTAATTCAAAGGTTGAGAA	TTCCCACGATATATTACTATTTAGTCTAAG	CGTCATTCTCAACCTTTGAATT	1251	13
WIAF-2142	G/A	A	G	TTTATGAACTGTCCGGCCCC[G/A]AGTCACTCAGCGTTTGCGGG	GAACTGTCCGGCCCC	CGCAAACGCTGAGTGACT	1255	16
WIAF-2145	C/T	C	T	TTAGCACTCAGACTTCCCCA[C/T]TCCCTAACTTTTGTTAATTG	TTAGCACTCAGACTTCCCCA	CAGCAATTAACAAAAGTTAGGGA	1258	17
WIAF-2149	A/C	A	C	CATATACGCTTCTGTCATTT[A/C]AACAAACTTCCAGAGAAAAC	GAATTCATATACGCTTCTGTCATTT	CCAGTTTTCTCTGGAAGTTTGTT	1262	1
WIAF-2150	C/T	C	T	TACACAGCTCAATGGGTCAC[C/T]TGGAACAAACTTGCTTGACT	TACACAGCTCAATGGGTCAC	ATATAGTCAAGCAAGTTTGTTCCA	1263	1
WIAF-2152	G/A	A	G	AATAAAACAAGTGAGAGACC[G/A]TTTACTTACATCAGTTCGGT	AGGTTAAATAAAACAAGTGAGAGACC	GTCTATAAACCGAACTGATGTAAGTAAA	1265	7
WIAF-2155	A/C	A	C	GGGATAACTGCATCTTTTGC[A/C]CCTTCACAACTAGAAACGAC	GGGATAACTGCATCTTTTGC	TGAGTCGTTTCTAGTTGTGAAGG	1268	9
WIAF-2158	T/C	C	T	TAGCCATATTTAGGATGAGA[T/C]GGATTGAGAGGCATGAACCA	CATTAGTAGCCATATTTAGGATGAGA	GGTTCATGCCTCTCAATCC	1271	10
WIAF-2162	A/G	A	G	CTTCTCCAAGGGGAGAACAG[A/G]CTGGAACTGCGGCTCTGCAA	TTCTCCAAGGGGAGAACAG	CAGAGCCGCAGTTCCAG	1275	1
WIAF-2165	A/G	A	G	GTCTAGGCCTGAGAATATTC[A/G]TTTCTAACAAGTTCCCAGGT	GGAGTCTAGGCCTGAGAATATTC	CTCAGGGTCACCTGGGA	1278	9
WIAF-2167	G/A	A	G	GAACAAGGACGCTTTGAAGA[G/A]GTGGAATTACTGTGCAAGGA	GAACAAGGACGCTTTGAAGA	CTCCTTGCACAGTAATTCCAC	1280	2
WIAF-2173	C/G	C	G	GAATGGAGGGAAATAAATGA[C/G]TGGATGGTCGCTGCTTTTTA	GGAATGGAGGGAAATAAATGA	AAGCAGCGACCATCCA	1286	2
WIAF-2177	G/C	C	G	TCCAATGAGCACTGTATGTA[G/C]AGAAAAGGGAAGGAGCAGGA	GCCTCCAATGAGCACTGTAT	CCTGCTCCTTCCCTTTTCT	1290	6
WIAF-2180	C/T	C	T	TGAACATTCCCGTGGTCTCC[C/T]TCTGAAAGCCGATGACCATC	AACATTCCCGTGGTCTCC	TGGTCATCGGCTTTCAGA	1293	X
WIAF-2182	G/A	A	G	GGCTCTAAGTCTTTATTGGG[G/A]AGAATACCCACCCACCTTCC	GGGCTCTAAGTCTTTATTGGG	GGAAGGTGGGTGGGTATTCT	1295	1
WIAF-2184	C/G	C	G	AAATGTTCCGGACCTAGATA[C/G]TGACGAAGGTAGCACGACAC	GAAAATGTTCCGGACCTAGAT	TGTCGTGCTACCTTCGTCA	1297	19
WIAF-2186	A/G	A	G	CAACAAAAGCTCTTCTGCAC[A/G]ATGGGAGGGAGACACCATTG	CCAACAAAAGCTCTTCTGCA	TGGTGTCTCCCTCCCAT	1299	6
WIAF-2189	G/A	A	G	AGTGCCGATGGTTACACAAT[G/A]TTGTAAATGTATTTAATCCC	AGTGCCGATGGTTACACAAT	TTCGTAAGTGGGATTAAATACATTT	1300	7
WIAF-2190	G/C	C	G	TCCTCTGGGGAGCAGGTCTA[G/C]GGCACGGAGGATGCAGGGCT	CCTCTGGGGAGCAGGTCTA	TGCATCCTCCGTGCC	1301	9
WIAF-2194	A/G	A	G	GACTCCCAGGAAAAGGTCCT[A/G]TCTTAGCTTCCTCCTCCCTA	ACTCCCAGGAAAAGGTCCT	AGTAGGGAGGAGGAAGCTAAGA	1305	3
WIAF-2195	G/A	A	G	AAGGAGGATTCATGAGCCCC[G/A]GTGACACAGATGGGGGCCCT	GGAGGATTCATGAGCCCC	GCCCCCATCTGTGTCAC	1306	11
WIAF-2196	C/T	C	T	AGTAGATGAAAAGGAAAGTA[C/T]AGAGAGGGCATTCAGGCCAA	TTAGAGGAGTAGATGAAAAGGAAAGT	GGCCTGAATGCCCTCTCT	1307	4
WIAF-2197	A/G	A	G	TCAGTAGGAGTGCAAGGGCT[A/G]TACCCCCGGAGCTAGACAGC	CAGTAGGAGTGCAAGGGCT	TGTCTAGCTCCGGGGGTA	1308	3
WIAF-2206	A/G	A	G	TGATGATGAATGGCGCGCTC[A/G]TACTCTTTACGGTCTTACAC	ATGAATGGCGCGCTC	GCATAAAAGTGTAAGACCGTAAAGA	1316	19
WIAF-2207	C/T	C	T	GGATGCTTCTTGCCAGGCTC[C/T]TTAAATTGTGCTGTAACCTG	TGCTTCTTGCCAGGCTC	CCCAGGTTACAGCACAATTT	1317	13
WIAF-221	T/C	C	T	CAGAAACTCTTTTTGGGCTC[T/C]TTTTTCTCCCTCTGGATCAA	ACAGAAACTCTTTTTGGGCTC	CTACTTTGATCCAGAGGGAGAA	2586	1

续表

WIAF 命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi 编号	NwLib chrom
WIAF-2214	G/A	A	G	CAAACATCCACATGGCACAA[G/A]CAGGGCCGGCCACTCCAGGC	AAACATCCACATGGCACAA	GAGTGGCCGGCCCTG	1324	21
WIAF-2216	T/C	C	T	ATAAGAATTGTTCTTCTGTT[T/C]GACAGTTGAAGTGGGTGTGA	TCACTGTTAATAAGAATTGTTCTTCTGT	CTCACACCCACTTCAACTGTC	1326	8
WIAF-2218	T/G	G	T	AGGGGAGACTTGGTATTTTG[T/G]TCAATCATTAAGAAGACAAA	GAGGGGAGACTTGGTATTTTG	CCCTTTGTCTTCTTAATGATTGA	1328	14
WIAF-222	T/C	C	T	AGCTTTGATTCTCCCCTTTT[T/C]TTGCATAAAGGCTGGGAAGG	AAGCTTTGATTCTCCCCTTT	TTCCCAGCCTTTATGCAA	2587	18
WIAF-2220	C/T	C	T	ACCTCAAAATGCAGAAACCA[C/T]TACAGATTAAAAGAGAAACA	CCTCAAAATGCAGAAACCA	GAGTTTCTCAAAGTGTGTGTGTG	1330	11
WIAF-2221	A/G	A	G	CTTAACCATTCAAACACCGC[A/G]TGACAACGAACCCAGTGGAC	TTAACCATTCAAACACCGC	TCCACTGGGTTCGTTGTC	1331	9
WIAF-2225	C/A	A	C	TGCAGAGATGATTCTTCTCA[C/A]CCCTTCTCTCAGGGTCGTGG	GGTGCAGAGATGATTCTTCTCA	ACGACCCTGAGAGAAGGG	1335	17
WIAF-2226	T/C	C	T	GGAAAGCTAAACGAACAAAA[T/C]GGTTTTAGTTTTGCTGAAGA	CTAGGAAAGCTAAACGAACAAAA	CAGTCTTCAGCAAAACTAAAACC	1336	8
WIAF-2227	C/T	C	T	AGTTCAGGAACTGCCTTTTA[C/T]GTCTGCAAATAAACTCCCAA	GAGTTCAGGAACTGCCTT	TTGGGAGTTTATTTGC	19263	15
WIAF-2228	G/A	A	G	ATTTTTGTTCAATGCCCACC[G/A]AGACATATAGAATTGGGAAC	AAACAATTTTTGTTCAATG	ATGTATCAGTTCCCAATTCTATAT	19445	8
WIAF-223	G/C	C	G	CTTCCTATTTTCTTTTGCTT[G/C]TTTTTTCTTTCACCTCAGAA	AATAAAACTTCCTATTTTCTTTTGCTT	GGGTTCTGAGGTGAAAGAAAAA	19485	2
WIAF-2233	A/G	A	G	TATTCATTCAATGAAGATTA[A/G]CACTGAGATCCAGAGAGGCT	CACTATTCATTCAATGAAGA	AGCCTCTCTGGATCTC	19402	14
WIAF-224	C/T	C	T	TAGAATTGTATAGCGATGAC[C/T]TCTTAATTATAATTTGATTT	ACTGAATAGAATTGTATAGCGATGAC	TTGCAGTCAAATCAAATTATAATTAAG	2588	6
WIAF-2264	T/C	C	T	AAATCCAGCTGTAACTTTTT[T/C]GGACTTGTCTTTTATTTCTT	GCAAATCCAGCTGTAACTTTTT	GCTTTTACTAAGCAAGAAATAAAAGAC	2286	7
WIAF-2265	T/C	C	T	TCAGCTCAACACTGCCTCCT[T/C]AGTGAGCCTGTGAACCACCC	CAGCTCAACACTGCCTCCT	GGTGGTTCACAGGCTCACT	2287	2
WIAF-2268	A/G	A	G	TACAGTTTACTGGAAGTTGT[A/G]TGAACTTGAGCAAGTGTCTC	GGCTGTACAGTTTACTGGAAGTTG	TTAAGAGACACTTGCTCAAGTTCA	2289	8
WIAF-2275	A/G	A	G	ACTGTGTGTTTGCTTGATTT[A/G]GGAGATAAAACCTGATCTCT	CCACTGTGTGTTTGCTTGATTT	TTTCTTAGAGATCAGGTTTTATCTCC	2292	9
WIAF-2278	A/C	A	C	TGGCTCAAAGGAAACTGTAG[A/C]AAATTCTTTTTTTTTATTTT	TTGGCTCAAAGGAAACTGTAG	CAAACTCCACTCTTTGAGTTAAAAA	2294	X
WIAF-228	C/G	C	G	GAGGTGACACAATCAAGACT[C/G]ACAGTAGCCTCAACCTCCTA	AGTGAGGTGACACAATCAAGACT	CATAGGAGGTTGAGGCTACTGT	2592	11
WIAF-2286	A/G	A	G	AGGGTACAGAGTGCTGCCAA[A/G]CACCTTAGAAAAATTACATG	GGGTACAGAGTGCTGCCAA	CGTGTCATGTAATTTTTCTAAGGTG	3607	13
WIAF-2287	T/C	C	T	TCTGCTTGCATTTTCAAGAT[T/C]CAATATATATCCAGATTGTT	GTCTGCTTGCATTTTCAAGAT	GGGAAAACAATCTGGATATATATTG	2299	7
WIAF-2288	T/C	C	T	TGTAACAAGAATTGATCCTA[T/C]ACTGGGACTACAGCCATGGA	CACAAATGTAACAAGAATTGATCC	CCATGGCTGTAGTCCCAGT	2300	4
WIAF-2291	A/G	A	G	TAGCAACAATATCAACAGAA[A/G]GGCTATATTAGAAAATTCTA	CCTATAATTTAGCAACAATATCAACAGAA	TGCAGGTAGAATTTTCTAATATAGCC	2303	4
WIAF-2292	C/T	C	T	TGTGTTACTCTATTTTGTTC[C/T]AGCCACCTGTGGCATTTCCA	GCATGTCTGTGTTACTCTATTTTGTTC	AAATGCCACAGGTGGCT	2304	1
WIAF-2293	T/A	A	T	CCTGGAATGCTCTTTCCCTC[T/A]GAGCTTTGCTTGGCTTACTT	CTGGAATGCTCTTTCCCTC	AAAGTAAGCCAAGCAAAGCTC	2305	17
WIAF-2295	G/C	C	G	TCCATTATACAACAGAAAAG[G/C]GGGCTGGAAAAGAAAGGTCA	ACATACATATCCATTATACAACAGAAAAG	GACCTTTCTTTTCCAGCCC	2307	X
WIAF-2298	A/G	A	G	ATGCTAGAAAATTATGCCTA[A/G]CCAAGTAGACAACTTAAGCA	GCAATGCTAGAAAATTATGCCT	TTAGGTGCTTAAGTTGTCTACTTGG	2310	5
WIAF-2299	T/C	C	T	GTGCTAGAACTAATCCCTCA[T/C]GGAGAACGTGGAACCACTGA	GGGTGCTAGAACTAATCCCTCA	CAGTGGTTCCACGTTCTCC	2311	3
WIAF-230	A/C	A	C	GGTGAGCACATATGGGTGCC[A/C]GCCCGAGACAGCAGGATAAG	TGAGCACATATGGGTGCC	CCTGCTGTCTCGGGC	2594	22
WIAF-2301	T/G	G	T	AGTGGTGAGATGCTCTGAGT[T/G]CAAGGCTGCTGACATGGTCA	CAGTGGTGAGATGCTCTGAGT	CCATGTCAGCAGCCTTG	2313	14
WIAF-2302	T/C	C	T	AATAAATGAATGCCATAATC[T/C]CTGTGTTTTTTGTCCCCACC	GAGTGAATAAATGAATGCCATAATC	TGAGAGGTGGGGACAAAAA	2314	14
WIAF-2303	G/C	C	G	CATTGAGGAAGTGTTTAAAG[G/C]AGAGAGATGACCCATCCATT	GCCATTGAGGAAGTGTTTAAAG	GAATGGATGGGTCATCTCTCT	2315	10
WIAF-2305	C/A	A	C	GCAGAGTCTGGGGGAGAAGA[C/A]AACGAGATAAAGCATGGCAA	CAGAGTCTGGGGGAGAAGA	TTGCCATGCTTTATCTCGTT	2317	4
WIAF-2308	C/T	C	T	AGTTGTAGCATTCAGAAGTC[C/T]CTCTTAGAGGTAGTTGTGCT	CCAAGTTGTAGCATTCAGAAGTC	ACGAGCACAACTACCTCTAAGAG	2319	7
WIAF-2311	A/G	A	G	AAGCCAGACAACACGAAAGT[A/G]TATAAAGAAAACAGTTAGTA	AAGCCAGACAACACGAAAGT	GGTGAAAGATTACTAACTGTTTTCTTT	2322	11
WIAF-2312	C/T	C	T	TCTTAGGCCCATCAGAGAAT[C/T]GAAGTCATGGGGAAAATTGA	TTCTTAGGCCCATCAGAGAA	TCAATTTTCCCCATGACTTC	2323	8
WIAF-2313	G/A	A	G	TGCACTATGGAACACCACAC[G/A]CAACTGAATGCAGATTCCAT	TGCACTATGGAACACCACAC	AATATGGAATCTGCATTCAGTTG	2324	11
WIAF-2318	A/T	A	T	GAGGTCTTAGTCATTGCATG[A/T]TGTATAACAATATTGTCACT	TTGAGGTCTTAGTCATTGCATG	TGAGTTCCTATTAAGTGACAATATTGTT	2328	2
WIAF-2319	A/G	A	G	CCTCCTTCACAACCTCACCA[A/G]ACTTGGCTTACCGGGAAGCA	CTCCTTCACAACCTCACCA	CTTCCCGGTAAGCCAAGT	2329	X
WIAF-2320	G/A	A	G	AATTATATTAACATGGTACA[G/A]ACAACTTCAGTTTAACATTG	GCATTGAATTAACTATAGATGTGTTAAGTA	CACTAGCAATGTTAAACTGAAGTTG	2330	12
WIAF-2321	G/C	C	G	TTCTCCCTTGTTTTCTTTTG[G/C]ATTGAAAAATACTGGTTTTC	TTTATTCTCCCTTGTTTTCTTTTG	ACTGTTAGAAAACCAGTATTTTTCAAT	2331	15
WIAF-2325	A/T	A	T	AATGCTATGTAGCAAATCTA[A/T]TCCCCTAAGCACAGTAATCA	TGATGTAATGCTATGTAGCAAATCT	TTGATTACTGTGCTTAGGGGA	2334	4
WIAF-233	G/T	G	T	AGATCCCCAAAGTGGTGGGG[G/T]CTTCCACGTGCAGTTGACTA	CCCCAAAGTGGTGGGG	TTAGTCAACTGCACGTGGAA	2595	9
WIAF-2333	G/A	A	G	AAAAGTAGAGATAATAATCA[G/A]TTCTTTACAACCGATGGTAA	GCCAGGAATAAAAGAAAAGTAGAG	TTAATTACCATCGGTTGTAAAGAA	3614	2
WIAF-234	T/C	C	T	ATCTCCAGGAGACTTTCATG[T/C]AGCCCAAAGTACAGCCTGGA	CATCTCCAGGAGACTTTCATG	CCAGGCTGTACTTTGGGCT	2596	1
WIAF-2343	C/T	C	T	AATTTCAGCATATGTATTAT[C/T]TGAACTAAATTTACAAAAGT	GGATATAGGAATTTCAGCATATGTATTA	TGTTCCACTTTTGTAAATTTAGTTCA	19324	16
WIAF-2351	A/G	A	G	ATCAAAGAAAGAATCTGAAT[A/G]TGAGGGAACTGCAGAAATTA	TGTTATAACATCAAAGAAAGAATCTGAA	TTAATTTCTGCAGTTCCCTCA	3626	18
WIAF-2359	C/T	C	T	AGTTACCCCACTAATACAAC[C/T]GAGAACCACTGACTTCAAAT	TGATGTAGTTACCCCACTAATACAAC	CATAATATTTGAAGTCAGTGGTTCTC	3631	12
WIAF-2362	A/G	A	G	TCCCTTATGTTGGCATTGCA[A/G]AGACACTGCACTTATCTGAG	CCCTTATGTTGGCATTGCA	AACCTCAGATAAGTGCAGTGTCT	3634	X
WIAF-2363	G/A	A	G	AAGGGCTATGTAACTACACA[G/A]TATGCACACCACAGCCATGT	TGGAAGGGCTATGTAACTACACA	CATGGCTGTGGTGTGCATA	3635	12
WIAF-2364	A/G	A	G	TTGAACCATGTGTAGACTGC[A/G]GGCACTTTAGAAAGAAGCTG	CTTTGAACCATGTGTAGACTGC	CTCAGCTTCTTTCTAAAGTGCC	3636	7
WIAF-2365	T/C	C	T	CATGTACGCACTAAAAAAAA[T/C]GTGTGCTTGCTGCTGCTGTG	CCAGTATCATGTACGCACTAAAAA	AGCAGCAGCAAGCACAC	3637	18

续表

WIAF 命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi 编号	NwLib chrom
WIAF-2372	T/C	C	T	TAATTAATGGTGGGCACATT[T/C]GCATGTGCTTACTGGGTCAT	CATATAATTAATGGTGGGCACA	TGACCCAGTAAGCACATGC	3644	9
WIAF-2374	T/C	C	T	AGACTGAGGTAAATAGTATT[T/C]ACGGCTGGAAATCAACATGC	GAGCAGTAGAGACTGAGGTAAATAGTATT	CATGTTGATTTCCAGCCGT	3646	8
WIAF-2375	A/C	A	C	CTCCCCCTTGTCCCTAGTTT[A/C]TAATTTCTCAGTGGACAAAT	TCCCCCTTGTCCCTAGTTT	TCCATTTGTCCACTGAGAAATT	3647	17
WIAF-2376	C/T	C	T	GCAGCCATTAACACCAGCAC[C/T]GATGCCACTTCTGTATCAGG	CAGCCATTAACACCAGCAC	TCCTGATACAGAAGTGGCATC	3648	2
WIAF-2377	T/A	A	T	CTTTGCCAGCAAAGCAAATA[T/A]CCGACTGACTGCTCCTTAGT	TTTGCCAGCAAAGCAAATA	ACTAAGGAGCAGTCAGTCGG	3649	8
WIAF-2382	A/C	A	C	CAGAGGAATAACTGATGTTC[A/C]CAATACCCCGACCCCTGACC	GGCAGAGGAATAACTGATGTTC	CAGGGGTCGGGGTATTG	3654	19
WIAF-2385	A/G	A	G	TATGCATTATCTTCACATGA[A/G]AAGGTTTCAGTTTATAAATG	GGTAAAATATGCATTATCTTCACATG	GTATTTAAGCATTTATAAACTGAAACCTT	3657	13
WIAF-2399	C/G	C	G	TAATGACTTTCTTGAACAAA[C/G]TGATTACGAAAGTGAAAGGC	ACAAAGTAATGACTTTCTTGAACAAA	GTAGCCTTTCACTTTCGTAATCA	3671	16
WIAF-2402	T/C	C	T	ACTGAATTCCTAAAATCTAT[T/C]ACACTGAGAGGAAAATGGAA	TCAGAATCTGACTGAATTCCTAAAA	AAAGCTTTATGCAAACATTTTCTT	3674	8
WIAF-2405	C/T	C	T	CTCCAAATACGAACAGTGCA[C/T]GCTGATGGCCTGCAGTCCTC	CTCCAAATACGAACAGTGCA	GACTGCAGGCCATCAGC	3677	17
WIAF-2407	G/A	A	G	CGTGCTTGGCTCTCTGGACG[G/A]TTCATTCTACATGGCTGCTG	GCTTGGCTCTCTGGACG	CAGCAGCCATGTAGAATGAA	3679	17
WIAF-2413	A/G	A	G	ACCATCTGTAACCCGAGCCC[A/G]CAGTGACCGGGACTTGCTGC	CATCTGTAACCCGAGCCC	GCAAGTCCCGGTCACTG	3685	17
WIAF-2420	A/G	A	G	TCAAGTCCCAGGCCCCAGGC[A/G]TCTTTCCTGCCCTGCCTTGC	GTCCCAGGCCCCAGG	GGCAGGGCAGGAAAGA	3692	12
WIAF-2422	C/G	C	G	GACTATTCTCTGAAGACAAC[C/G]AAGAGAAAATTGCAAAAAGA	GCAAATGACTATTCTCTGAAGACAA	TTGTCTTTTTGCAATTTTCTCTT	3694	
WIAF-2423	G/C	C	G	AGTGCTCAGCTACTTCTCCT[G/C]CACTTTGAAAGACCCCTCCC	CAAGTGCTCAGCTACTTCTCCT	GGAGGGGTCTTTCAAAGTG	3695	12
WIAF-2426	T/G	G	T	CTTGATTGTATATAAGATAA[T/G]TCATACTGGAGAAAACTCCC	CGGTTGTCACACTTGATTGTA	TCTGGGAGTTTTCTCCAGTATG	3698	
WIAF-243	C/T	C	T	GTACTTGCTCTGTGCCGTAT[C/T]TGCTCCAATCACCCATTCCA	GGTACTTGCTCTGTGCCGTA	GAATGGGTGATTGGAGCA	2115	15
WIAF-2434	C/T	C	T	AAATATGGACTATCTTCAAA[C/T]TGCACAAATGATGCATGAAT	AAATAGGAAATATGGACTATCTTCAAA	TTCATGCATCATTTGTGCA	3706	
WIAF-2436	A/G	A	G	AAGGATTGTAATTTGGGTAA[A/G]CTGAGTCACGGTGGCCCTGA	AACAAAGGATTGTAATTTGGGTAA	GGGCCACCGTGACTCA	3708	21
WIAF-2438	C/G	C	G	ATCCCAGCCCTTAGCATCAA[C/G]TCATCTTCAGTCTTTGAGTC	CCCAGCCCTTAGCATCAA	AAGGACTCAAAGACTGAAGATGA	3710	10
WIAF-2441	C/T	C	T	ATAAGGGAAAATGGGAAGAA[C/T]AGAGTGAAATTAAAGCAAAT	GAATAAGGGAAAATGGGAAGAA	CCAAGATTTGCTTTAATTTCACTC	3713	1
WIAF-2445	G/A	A	G	TACTACTGTGACAATTTAGC[G/A]ATCCTTCAAATGGGAAAATT	ATACTTCCTACTACTGTGACAATTTAGC	GAATTTTCCCATTTGAAGGAT	3717	17
WIAF-2449	G/A	A	G	GGAGGGACAGAGAAACATGA[G/A]CTGGGGAGTAGGCTCTGACA	GGAGGGACAGAGAAACATGA	GTCAGAGCCTACTCCCCAG	3721	2
WIAF-2450	T/C	C	T	TCAGCTGCAACTTCTGCTTT[T/C]CTCCTTTGCCTCTGCTCCTT	TCAGCTGCAACTTCTGCTTT	GGAGCAGAGGCAAAGGAG	3722	17
WIAF-246	G/A	A	G	ACACTGAATAAACATCTCAC[G/A]AACTGTCGCTCCTAGATTAC	GCTCTTTAACACTGAATAAACATCTCA	TTTGTAATCTAGGAGCGACAGTT	2602	14
WIAF-2460	C/T	C	T	AACCTTAGATACATAGCCGA[C/T]TGTATACAGAGGTTCATCTC	TGAAAACCTTAGATACATAGCCGA	GCCCCAAAAGTCAATAGTGTT	3732	X
WIAF-2467	G/T	G	T	GCTGTAAGCCAGCATGGGGT[G/T]GGGGAGGTGATTATGGCTGG	TGTAAGCCAGCATGGGGT	AGCCATAATCACCTCCCC	3739	12
WIAF-2477	A/G	A	G	CAGCTATTGTTATTTCAAAT[A/G]TATCTTCTGCTCCCTTTTCT	GCTGTCAGCTATTGTTATTTCAAA	GGAGAAAAGGGAGCAGAAGA	3749	21
WIAF-2482	C/T	C	T	TCCTTACTTCCCCCATAGAT[C/T]CCTGACAATGTGCTGCAGAA	CTTCCTTACTTCCCCCATAGA	CTGCAGCACATTGTCAGG	3754	5
WIAF-2488	C/T	C	T	ACCAGCCCTGGCCCTGCCCT[C/T]AGCTGCATGCCACCCTCATA	CAGCCCTGGCCCTGC	GGGTGGCATGCAGCT	3760	11
WIAF-2489	A/G	A	G	AGCATTAGGGATGAGTTCTC[A/G]GAAGTGATTCTGAACTGAGC	GCAGCATTAGGGATGAGTTCTC	CGTGCTCAGTTCAGAATCACTT	3761	15
WIAF-249	C/G	C	G	TGAGACTGAAGGGAAATCCC[C/G]CTTTCTTTCTACCAGCCCTG	TGAGACTGAAGGGAAATCCC	GCAGGGCTGGTAGAAAGAAA	2604	3
WIAF-2490	T/C	C	T	TCCGAATACAGTACTTCTTT[T/C]GAAAAAATACACAATGGGAA	GGCATCCGAATACAGTACTTCTTT	GTCAGTTCCCATTGTGTATTTTT	3762	5
WIAF-250	C/T	C	T	TACAGAGACAAAACGGCCTC[C/T]GGCTCTCAGAGCATAATCCT	ACAGAGACAAAACGGCCTC	AAGGATTATGCTCTGAGAGCC	2605	3
WIAF-2501	A/C	A	C	TTGATTGAGCAATCTAGGGG[A/C]TATGTGACAGGGGTTTCATG	TTGATTGAGCAATCTAGGGG	GTTGTGTTCTGTACCAGTGCA	3773	5
WIAF-2513	T/G	G	T	GGTCTTTTATCCCTCGCCCC[T/G]CTCCCACTTTTCCCCTCAAG	CCATACTCGGTCTTTTATCCC	TGAGGGGAAAAGTGGGAG	3785	21
WIAF-2514	A/C	A	C	TTGCTTAAGCCACTTGGGTA[A/C]TCCATTCCAGCTCTGCACCT	TTGCTTAAGCCACTTGGGTA	TGCAGAGCTGGAATGGA	3786	2
WIAF-2515	C/G	C	G	ATGTTGCAGTGCGGATGAAT[C/G]TTATCATGATGCTAAGTGAA	GTTGCAGTGCGGATGAA	ACAGGTAGAATGATACATTTGGATT	3787	19
WIAF-2525	G/A	A	G	GTATTTACTTGAGGGCAACA[G/A]AATTACGGCTTAACAACACA	TGGTATTTACTTGAGGGCAACA	TTTAGTGTGTTGTTAAGCCGTAAT	3796	8
WIAF-2528	G/A	A	G	CAACTGGAAAGCTTTTACAC[G/A]TGCTTCAGAATGCGGCAGTA	CCCAACTGGAAAGCTTTTACA	GCCGCATTCTGAAGCA	3799	5
WIAF-2531	G/A	A	G	AATTATGGACTTTAAAGCTC[G/A]ACATAAAATTAGTAGCTTCA	AGATAACAATTATGGACTTTAAAGCTC	CCCTTTTGAAGCTACTAATTTTATGT	3802	12
WIAF-2534	G/A	A	G	ACATTTTGAATATAGCTATC[G/A]TTTTAACAAACCTCATTATG	TCCAAATAAACATATAGACATTTTGAA	AGTGATCATAATGAGGTTTGTTAAAA	3805	13
WIAF-2538	A/C	A	C	CTGAATTAGTCAAGCAGGTC[A/C]GATACTATTGTCTGCTAGAT	GGCTGAATTAGTCAAGCAGGTC	TTAAAAGTATTACAGAAGCAAACTTTTT	1340	22
WIAF-2541	T/C	C	T	CAAAGCAGCTAAAGGAATAT[T/C]TACACCACCCACCCCCTTTT	TGCAAAGCAGCTAAAGGAAT	GGGGGTGGGTGGTGTA	1343	17
WIAF-2544	G/A	A	G	TGAGATATTTCTCACAATCA[G/A]ACAAGAGCCAGGGCCCCCAG	CCTGCTGAGATATTTCTCACAATC	GGGCCCTGGCTCTTGT	1345	22
WIAF-2546	A/C	A	C	GGAGAAATGAAGATTTAATC[A/C]TTCCTAAATTTAAGTCAGAT	CATGGGAGAAATGAAGATTTAATC	GATTTCTTTCTATCTGACTTAAATTTAGG	2060	6
WIAF-2555	G/A	A	G	CGTCTATCTGGCGAGGAGGC[G/A]CGGCATGGAGTCCAGGGTGG	TCTATCTGGCGAGGAGGC	CCTGGACTCCATGCCG	2068	13
WIAF-2560	T/C	C	T	GAACTGGCCTGGACGGGATT[T/C]GAGGGCAGCTGGCGGGGGCT	TGGCCTGGACGGGATT	CCGCCAGCTGCCCTC	3808	17
WIAF-2563	G/C	C	G	GAGCACAAGCCAAAGATACA[G/C]AGAAATAAAGAGTGTACAAT	AGAGCACAAGCCAAAGATACA	ACTGTGCTCTTATTGTACACTCTTTATT	3811	3
WIAF-2567	C/T	C	T	TGTCTAAATTCCACTAAAGC[C/T]GCTTCTCAGTGGTCTCCAGT	GCTTCTGTCTAAATTCCACTAAAGC	CACTGGAGACCACTGAGAAGC	2341	9
WIAF-2569	T/C	C	T	TGCATTTTCCCATCGTTCTT[T/C]AGATTTAGGAGAAACCCACG	GCATTTTCCCATCGTTCTT	CCGTGGGTTTCTCCTAAATC	3815	11
WIAF-2572	T/C	C	T	TAAAGTTCAAAGTGGTAATG[T/C]GCTTTCTCCAAGTACCTATG	AGAAGAATCTAAAGTTCAAAGTGGTAAT	CTTAACATAGGTACTTGGAGAAAGC	2342	3
WIAF-2573	G/C	C	G	TGGCGCCACAGAGGAAGCTG[G/C]CTGCCGGAGGACCCCCACTT	CGCCACAGAGGAAGCTG	GGGGTCCTCCGGCAG	3818	17
WIAF-2574	C/G	C	G	CTCGGCTTGAAGATCCCTAA[C/G]CTGCATAGGTTTTTGGAATT	TCGGCTTGAAGATCCCTAA	TTTAATTCCAAAAACCTATGCAG	2343	19
WIAF-2576	C/G	C	G	TCAACTGAAATGAAAATATA[C/G]AAAGGAGGATGGGTAAAATT	CTCCTTTCAACTGAAATGAAAATAT	TTTAATTTTACCCATCCTCCTTT	2345	15
WIAF-258	A/C	A	C	TCCTCATGGGTTTGTGACAG[A/C]CCCTGCGTGCTCAGTGCCCT	CCTCATGGGTTTGTGACAG	CACTGAGCACGCAGGG	2612	6

续表

WIAF 命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi 编号	NwLib chrom
WIAF-2581	A/C	A	C	CATAACCACCACAGTCAGAG[A/C]TACTCTGCTCCACCATCACA	GCATAACCACCACAGTCAGA	TGTGATGGTGGAGCAGAGTA	2347	10
WIAF-2588	T/C	C	T	AACTGTGCCCATACTGATAG[T/C]GCATGGCAGGAGAATTTTGA	GCAACTGTGCCCATACTGAT	CAAAATTCTCCTGCCATGC	2351	10
WIAF-2589	G/A	A	G	GCAGTTGTTACCCATTAAAT[G/A]ATCAATTTCTATCCCTCAAT	TCAGCAGTTGTTACCCATTAAA	TGAAGACTGTATTGTATTGAGGGAT	3824	X
WIAF-2590	A/G	A	G	CTTACTGTGTCATCCTATCC[A/G]TTCCCTTCCCTGAGCCTGGG	TCTGTACTTACTGTGTCATCCTATCC	AGGCTCAGGGAAGGGAA	3825	2
WIAF-2592	C/T	C	T	GTGCAAATGCCTTAAGATAG[C/T]GTAAGTCTAGGGAGTGTGAG	CAAGTGCAAATGCCTTAAGAT	CAAATCTCACACTCCCTAGACTTAC	2352	4
WIAF-2597	C/T	C	T	CAATCACACTAAATAGAAAC[C/T]TTGTGCCATGTACAACCACT	GTACAATCATCAATCAAATTATAGAACA	AGTGGTTGTACATGGCACAA	3828	6
WIAF-2601	G/C	C	G	TCCTTCCAGAGGCTGAGGGA[G/C]AGAATGCGTTTTCTTGCCTT	TTCCAGAGGCTGAGGGA	CGCTTCTGGAAAAGCTGG	3832	7
WIAF-2612	C/T	C	T	TCAGCCTGCACCCTCTTCTC[C/T]TCTGGCTTCCCTCTCTGAAA	AGCCTGCACCCTCTTCTC	TTCAGAGAGGGAAGCCAGA	3839	3
WIAF-2616	T/G	G	T	AGGTTGGACTCTGCAGAAAC[T/G]GTGACTCAGTGTTGTTTGGG	CAGGTTGGACTCTGCAGAAA	ACCCAAACAACACTGAGTCAC	2362	7
WIAF-2618	C/A	A	C	TCCTCTTTCTCTAGGCTGAC[C/A]CAGGTCGTTTTACTGACCGA	TCTTCCTCTTTCTCTAGGCTGA	TCGGTCAGTAAAACGACCTG	2363	16
WIAF-2626	C/T	C	T	CACTGTGAATTAGAGCCAGA[C/T]GCTTAAGTCCAGGGTGAGAC	GCACTGTGAATTAGAGCCAGA	TGTCTCACCCTGGACTTAAGC	3847	3
WIAF-2627	A/G	A	G	CAGAACCACTTGCCATCAAC[A/G]GCTGGGGAATCCTAGCCAGT	CAGAACCACTTGCCATCAAC	GGCTAGGATTCCCCAGC	2366	6
WIAF-263	A/C	A	C	CCCTTCCTCCCTGTCTCTGC[A/C]TCTGACTCTCTTTTGAGGTC	CCTTCCTCCCTGTCTCTGC	GGGACCTCAAAAGAGAGTCAGA	2616	7
WIAF-2630	C/G	C	G	GCAAGACCAGAACACTCTGT[C/G]CCTATGGTAGATCTACAAAA	GGCAAGACCAGAACACTCTG	GGCCTCTTTTGTAGATCTACCAT	2369	15
WIAF-2633	A/T	A	T	CTGCTGGGAATCAAGCATAA[A/T]CACATCCATCTGATACAAAG	TGCTGGGAATCAAGCATAA	AAGATACTTTGTATCAGATGGATGTG	2370	8
WIAF-2635	C/G	C	G	CTAGTGATAACTATACAAAT[C/G]TGGACATGGATAATGAACAG	GGGTAAAATATTTATTTCCATCTTCA	CCTGTTCATTATCCATGTCCA	3849	2
WIAF-2636	G/A	A	G	TCGAGTCTCATGTTCCAACC[G/A]ACCGTCTGGCTCAGTGACTG	TCGAGTCTCATGTTCCAACC	GTCACTGAGCCAGACGGT	3850	1
WIAF-2637	C/T	C	T	ATGTTGTGTGAGTTTTAACA[C/T]ATTTGGGCTTGCATTTGTTC	GCAACCTTTAATATATATCATGTTGTG	GCTGATAAACTTTTCATATTACACATTG	3851	19
WIAF-2638	A/G	A	G	GTCCTCATCAGTGAAACCAG[A/G]GTAAGAGAAACACCAAAGGC	GGTCCTCATCAGTGAAACCAG	TGCCTTTGGTGTTTCTCTTAC	2372	2
WIAF-2639	T/A	A	T	TTTATCTCCCATATTTCAAA[T/A]ACATGCTGTCAAGTTTGGCT	AGGAATTTATCTCCCATATTTCAA	AGCCAAACTTGACAGCATG	3852	19
WIAF-2640	A/C	A	C	TAAAAATAGCTCCAGATGAC[A/C]GAACAGTATTGCATACACAG	CATCTTAAAAATAGCTCCAGATGAC	GGACCTGTGTATGCAATACTGTTC	3853	10
WIAF-2641	C/A	A	C	CAGTGGACACAAATCCCTCC[C/A]TCATTATCCAGGCATGGATG	GCAGTGGACACAAATCCC	CATCCATGCCTGGATAATGA	3854	19
WIAF-2642	A/G	A	G	GAGTACCAGTCGCTAAATGT[A/G]CAGTAAAAGAAGCAGGCAAC	CGAGTACCAGTCGCTAAATGT	TGTTGCCTGCTTCTTTTACTG	2373	18
WIAF-2643	G/C	C	G	GGAGAAAGACAGAATAAATA[G/C]AACAGTAGAATGATCACAGA	TGGGAGGAGAAAGACAGAATAA	CCCATCTGTGATCATTCTACTGTT	3855	21
WIAF-2644	T/A	A	T	CTTTTCAGCATTCAGTTTTG[T/A]GGCTCATGGGAAATGTTACT	GCTTTTCAGCATTCAGTTTTG	AAGTAACATTTCCCATGAGCC	3856	1
WIAF-2649	T/C	C	T	TTCCGGGGACCACGATCGGC[T/C]GCTTGTCTGCTACCCAGCAT	GGGACCACGATCGGC	GCTGGGTAGCAGACAAGC	19486	2
WIAF-2650	A/G	A	G	TGTTGGGTGGAACAGCAGAG[A/G]TTAATAGGTTTCCACAAGCC	TTGGGTGGAACAGCAGAG	TGGCTTGTGGAAACCTATTAA	3860	16
WIAF-2651	C/G	C	G	TGCTCCCACAGGGGCCATGT[C/G]CTCTCTTAATAGGTGGCACT	CCCACAGGGGCCATGT	GGGTAGTGCCACCTATTAAGAGA	3861	7
WIAF-2652	A/C	A	C	CTTGGGCTTGATCTTGGCTG[A/C]GAGGTGCGGAATGTCATCGT	GGGCTTGATCTTGGCTG	GATGACATTCCGCACCTC	2375	16
WIAF-2654	A/G	A	G	GTATTGCAATCCCAGTTGAA[A/G]GTCCTGGGGAGGTGTTTGTA	GGTATTGCAATCCCAGTTGAA	CAAACACCTCCCCAGGAC	2376	17
WIAF-2658	G/A	A	G	CATACTATAACTTCTATTCT[G/A]ACTTAAACTCTGTTCTATTC	TGTCAAAGCCATACTATAACTTCTATTC	TGACACAAGAATAGAACAGAGTTTAAG	2377	12
WIAF-266	A/C	A	C	AGGCCAGGCAATTTTATTTG[A/C]GCCCTAGGAGGGTTACTATA	GGCCAGGCAATTTTATTTG	AAATTATAGTAACCCTCCTAGGGC	19446	20
WIAF-2661	A/T	A	T	GTGCCCCAAAGATGCCTCAA[A/T]GCCTAGAAAAGTGCCTTCCC	CCCCAAAGATGCCTCAA	TGGGGAAGGCACTTTTCTAG	2378	12
WIAF-2662	T/C	C	T	GATGTTGCAGAATCTCAGAG[T/C]TGGATACTTGCAGGAAGCTG	GGATGTTGCAGAATCTCAGAG	CAGCTTCCTGCAAGTATCCA	3867	19
WIAF-2665	A/G	A	G	GGAGGAGGGAGGCAGCCAAA[A/G]GTGAATAGCACAAGATGAGC	GAGGGAGGCAGCCAAA	CTGCTCATCTTGTGCTATTCAC	2380	3
WIAF-2667	G/A	A	G	CAATAGAGTTTATAGTCTTA[G/A]GGGCAACAATAGGGAATGAG	AAAGGCTGATGTGCTAACACA	TCATTCCCTATTGTTGCCC	2381	1
WIAF-2669	T/G	G	T	CAATAAAAAAAAACCAACAC[T/G]TTGCTATGTGGTCGTTCCAA	GTTTACCACAGAGTGTTACAATAAAAA	TGGAACGACCACATAGCAA	3870	X
WIAF-2670	C/T	C	T	TTTCATTAAGTGCTAAGTAT[C/T]GTCTGCAACTGATTTTCCTG	TCCTACATTTTCATTAAGTGCTAAGTA	TCAGGAAAATCAGTTGCAGAC	3871	2
WIAF-2671	C/T	C	T	AAGCAAATTATGTTTAAAAG[C/T]TGGGCTGCCATCCATTATCC	CACATTTCTGTGTAAGCAAATTATG	ATGGATGGCAGCCCA	3872	2
WIAF-2672	G/A	A	G	ACATTCTAAATACACTAAAC[G/A]TTATCTAATGAAGTCACACT	GCACAAATTTACACATTCTAAATACAC	GAAGACCAGTGTGACTTCATTAGA	3873	8
WIAF-2674	T/C	C	T	TAACTTCATGCAGAACATTC[T/C]GGCTGTGTTGGGAATGAGAA	CGTAACTTCATGCAGAACATTC	CTCATTCCCAACACAGCC	2383	1
WIAF-2676	A/C	A	C	TCAGAAGCACCAGGAAGGCG[A/C]ATAGGGATGCCAGAGCCCCA	AAGCACCAGGAAGGCG	GGGCTCTGGCATCCCTAT	3875	1
WIAF-2677	A/C	A	C	GACTAGGAGGAGAAAGGGGG[A/C]GTGGATTTATGATTTTATGG	ACTAGGAGGAGAAAGGGGG	TTGCCATAAAATCATAAATCCA	3876	14
WIAF-2679	G/A	A	G	CTTTTGTATTTGCCTTTAGA[G/A]GGAGGATAACAGTAACTTAT	TTGCTTTTGTATTTGCCTTTAGA	CTTCTACAAATAAGTTACTGTTATCCTCC	2385	12
WIAF-268	C/T	C	T	GTGGGACAAAAGGCTTGTCA[C/T]CTGTCAGAAACATTGAAAAC	GGGACAAAAGGCTTGTCA	GCTGTTTTCAATGTTTCTGACAG	19487	10
WIAF-2686	G/A	A	G	CCAAACAATGCCGTGGAAGA[G/A]GTGATCAAAGCAGAAGGGGC	AAACAATGCCGTGGAAGA	CCCCTTCTGCTTTGATCAC	19488	2
WIAF-2689	T/C	C	T	GAGAATGGCTATTGGTTCTC[T/C]TTTGTAGTCTCTCACATTTA	GGAGAATGGCTATTGGTTCTC	CTTGATGTGATAAATGTGAGAGACTAC	19266	12
WIAF-2690	G/C	C	G	AAGAAGTAGCAGAGCCTCCA[G/C]ATGCTTTTATCTTCCCCTAG	CTGAAAGAAGTAGCAGAGCCTC	GATCTAGGGGAAGATAAAAGCAT	19222	6
WIAF-2692	T/C	C	T	CGTTCCAGGTTTAAAGATTA[T/C]AGCTAAACACTTTGAGGGAC	TGAACGTTCCAGGTTTAAAGA	GCAGTCCCTCAAAGTGTTTAGC	19331	9
WIAF-2693	G/A	A	G	GATCATTCACCACTTTGCTG[G/A]GATTTTGCCTTAGTGCCACC	GGATCATTCACCACTTTGCTG	GGTGGCACTAAGGCAAAATC	19489	7
WIAF-2696	A/G	A	G	TTGCCAAGGAAGTAGTTGTC[A/G]AGATCCAGGTCAAAAGGAAA	TTTGCCAAGGAAGTAGTTGTC	CTTTCCTTTTGACCTGGATCT	19490	11
WIAF-2699	T/C	C	T	TGTGGATTTCCTACATGATG[T/C]CATGGAAAACACTGGTTTTG	CTAGTGTGGATTTCCTACATGATG	CCAAAACCAGTGTTTTCCATG	19491	

续表

WIAF 命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi 编号	NwLib chrom
WIAF-270	A/G	A	G	CCAGAAGACTCAGCTGCTTG[A/G]GGCATGTTCCCACCCTGGAC	CCAGAAGACTCAGCTGCTTG	AGGGTGGGAACATGCC	2619	1
WIAF-2701	T/C	C	T	AGTGGGGTCCTCCCCTGAAA[T/C]AGGTAAATGAGAACTGGGAG	GGGGTCCTCCCCTGAAA	CACTCCCAGTTCTCATTTACCT	19492	2
WIAF-2703	G/T	G	T	TGTTCCAATTAGTTTTAATG[G/T]TACTTTGAAAGGCCTCACAT	TGTATACATTTTCTAATATTGTTCCAATT	CATGTGAGGCCTTTCAAAGT	19173	3
WIAF-2705	G/A	A	G	GGACCTGCTCAAATGGATCT[G/A]AGACTCTAGGGGGGAAGCCA	GGACCTGCTCAAATGGATCT	TTCTTCTCTGGCTTCCCC	19223	5
WIAF-2706	A/G	A	G	GTGTGAATAAAGAAATGAAA[A/G]GCATTTAGAGAGAGAGGTGA	AGTGTGGTGTGAATAAAGAAATGA	CATTCACCTCTCTCTCTAAATGC	19259	13
WIAF-2709	C/T	C	T	TGCAGATGGAGAAAGTAAGG[C/T]TGGAGAGTAAAGGCAGAGCT	CATTTTGCAGATGGAGAAAGTA	CAGCTCTGCCTTTACTCTCCA	19493	6
WIAF-271	G/A	A	G	AGGCTGAGAGCTGGCACCAC[G/A]CACTGGTTTCTAAATCTCTG	GCTGAGAGCTGGCACCAC	AAGCCAGAGATTTAGAAACCAGT	2620	10
WIAF-2713	T/C	C	T	GCTCACATTTTTGAATCTTG[T/C]CCTCAATACCTGAACCTGAG	CTGCTCACATTTTTGAATCTTG	CCTCAGGTTCAGGTATTGAGG	19494	13
WIAF-2714	C/T	C	T	ATGAAGCTTTCCTAGGACCA[C/T]TAAACAATTGTAGTTCATTA	CATGAAGCTTTCCTAGGACCA	GGATCAGTAATGAACTACAATTGTTTA	19279	9
WIAF-2716	C/A	A	C	CTCAACAGAACAAGCCTGCT[C/A]CTGCCTCAGGGCCTCTGTCC	TCAACAGAACAAGCCTGCT	AGAGGCCCTGAGGCAG	19295	12
WIAF-2718	C/A	A	C	CCTGCCTCAGGGCCTCTGTC[C/A]TTATTCTTCCTCTTCCCAGG	GCCTCAGGGCCTCTGTC	CCCTGGGAAGAGGAAGAATAA	19296	12
WIAF-272	G/C	C	G	CTCTCCATACGTAGGTCCTG[G/C]TCTCCTATCACATTGCCACG	GTCTCTCCATACGTAGGTCCTG	CGTGGCAATGTGATAGGAGA	19197	10
WIAF-2721	G/A	A	G	ATGAGAGCAGTCCTGCCTGG[G/A]TGGGGCAAGACAGAAAGGCT	AGAGCAGTCCTGCCTGG	CCTTTCTGTCTTGCCCCA	19495	9
WIAF-2723	A/G	A	G	TTTGTATCAGAAAAACCTGC[A/G]TTCAAGTTCCAGCTCTCCCA	CCTTTGTATCAGAAAAACCTGC	GGGAGAGCTGGAACTTGAA	19176	5
WIAF-2724	T/C	C	T	AGACCCAAGATGTCAATACA[T/C]AGTGCTGCATTGAGAGACCA	CAGAATAGCCTCTGACCAAAAA	TGGTCTCTCAATGCAGCACT	19317	8
WIAF-2728	G/T	G	T	GTCTGGAGAGATTTGCTGTG[G/T]AATATATTTCTAAGGAAGAA	AGTCTGGAGAGATTTGCTGTG	CCTCTGGATTCTTCCTTAGAAATATATT	19496	7
WIAF-2729	T/G	G	T	TCAAGTCTGATAATGAAATA[T/G]ATGTGCAAAACTGACAGTGG	GGATACATTTTCAAGTCTGATAATGA	CCACTGTCAGTTTTGCACAT	19497	10
WIAF-2730	A/G	A	G	TTAGATATTTTCTGGAAGCC[A/G]ATGTCACGTTTTCCTTCCAC	CTGATTAGATATTTTCTGGAAGCC	GTGGAAGGAAAACGTGACAT	19310	3
WIAF-2735	C/T	C	T	TATTAGTTCCTTAGGGCTGC[C/T]GGAGCAGAGTACCTTAAGCT	AAAGTAACATTGTGTATTAGTTCCTTAGG	CCAGCTTAAGGTACTCTGCTCC	19307	4
WIAF-2736	G/C	C	G	AACAGACATTCATTCTCTCA[G/C]AGGTCTGCAGGCCAGAAGTC	CCTTAAGCTGGTGACTTAAAACAA	TTGTTTTAGACTTCTGGCCTG	19308	4
WIAF-2737	A/G	A	G	CTGCAGGGGAACCAACCTAG[A/G]AGCCTGAACTGGGGGCCTGG	GCAGGGGAACCAACCTAG	GCCCCCAGTTCAGGCT	19264	8
WIAF-2738	G/T	G	T	ATGATTCTAGATTACAGAGA[G/T]TTTTTTCAAACCAATGCACT	TGAAAGGTATGATTCTAGATTACAGAGA	TCTAGTGCATTGGTTTGAAAAA	19265	8
WIAF-2739	C/T	C	T	AGATGTATAATCACTTATCA[C/T]GTTTTTGCTTGCATCATGTG	CAAGCCTAGATGTATAATCACTTATCA	CACATGATGCAAGCAAAAAC	19498	2
WIAF-2741	G/A	A	G	AAGGAGCAGATTACAGCAAA[G/A]AACATTATCCACCCAGCACC	GAAGGAGCAGATTACAGCAAA	GGTGCTGGGTGGATAATGTT	19166	11
WIAF-2749	C/T	C	T	TTGTGGGTAGATAATAGATT[C/T]AACTCAAAATTCCTCCCCCA	AATTGAATCCTTGTGGGTAGATAA	GGGGGAGGAATTTTGAGTT	19292	3
WIAF-2750	G/A	A	G	CCAGTATGTGTAGTTTGCAG[G/A]CATCATCCTCCCAAAGCAGG	AACCCAGTATGTGTAGTTTGCAG	TGCTTTGGGAGGATGATG	19499	16
WIAF-2752	A/T	A	T	CAAAGCAGGAAGTAGAATAA[A/T]GACAGTAGATTTAGAGACTC	CCCAAAGCAGGAAGTAGAATAA	TCTCTGTGGAGTCTCTAAATCTACTG	19323	16
WIAF-2753	A/C	A	C	TCACAGGCAGAGAGGGATAC[A/C]CTTGCAGATCAGCTCCATGG	TCACAGGCAGAGAGGGATAC	CATGGAGCTGATCTGCAAG	19293	3
WIAF-2754	T/C	C	T	AGTCACTATCTCCCTGACTC[T/C]AGTTATAAGGTAAATTCTGT	TTCAGTCACTATCTCCCTGACTC	GTATTTTAGTCAACACAGAATTTACCTTA	19500	13
WIAF-2758	C/G	C	G	GGATCCAGGATCCTTTGCTT[C/G]TTTTTTGCCTCACTCCCCTC	GGATCCAGGATCCTTTGCTT	GGGGAGTGAGGCAAAAA	19154	11
WIAF-2759	A/G	A	G	CCCTCAGCAGTGTGCTTTGC[A/G]TGCTACAGGTCTTCCCACCT	CCTCAGCAGTGTGCTTTGC	GGTGGGAAGACCTGTAGCA	19212	9
WIAF-2760	C/T	C	T	TCTGATACCTCCACTTCAAC[C/T]TACGATGTGTAATTCCCAGT	TGAATGTTATTCTGATACCTCCAC	GTAAACTGGGAATTACACATCGT	19318	17
WIAF-2763	G/A	A	G	AGATTCCTGGGGTGTGCTAT[G/A]GAAGTACTTTGCCCAAATTT	GATTCCTGGGGTGTGCTAT	GAAAATTTGGGCAAAGTACTTC	19291	8
WIAF-2765	G/A	A	G	TAAATGGCGTGATATTACTA[G/A]GCATCACCAAATTGCAGAAT	GAAGCTAAATGGCGTGATATTAC	TTCTGCAATTTGGTGATGC	19501	4
WIAF-2766	A/G	A	G	AGTCCTTCCTAAAGTGTCTT[A/G]GCTTGCCACATGTACAGCAG	GGCAGTCCTTCCTAAAGTGTC	TGCTGTACATGTGGCAAGC	19230	6
WIAF-2767	A/G	A	G	AGAGCCTCTGCCTTTTTCTT[A/G]TATCTTCTTTTCCTGTCTTG	CATTAGAGCCTCTGCCTTTTTC	CCCAAGACAGGAAAAGAAGAT	19231	6
WIAF-2768	C/T	C	T	CTCCTGGTTATGATGGCAAG[C/T]GTGTGAATAGAGGGGTGGCA	TCCTGGTTATGATGGCAAG	GCCACCCCTCTATTCACAC	19198	11
WIAF-2769	A/G	A	G	ATCATGTCATTTAATCACCC[A/G]TGCCATCAGGCCTGTTTTAT	CGTATCATGTCATTTAATCACCC	AAACAGGCCTGATGGCA	19502	11
WIAF-2780	G/C	C	G	TGCTGCTGTCTATACCCACA[G/C]ATAGGGAACTGGATGCTTAG	TGCTGCTGTCTATACCCACA	CATCTAAGCATCCAGTTCCCT	19503	18
WIAF-2781	A/G	A	G	TCTCCAGTAAAAACTCAATG[A/G]ATGCTGTTGGAGTCACCACC	GATTTTCTCCAGTAAAAACTCAATG	GGTGGTGACTCCAACAGCAT	19206	13
WIAF-2782	T/C	C	T	CTATTCCATCACTTCTTACC[T/C]GGGCTACTGCAATAGCCTTG	GAACCTATTCCATCACTTCTTACC	GAATAATTACAAGGCTATTGCAGTAG	19169	9
WIAF-2783	G/A	A	G	ATATTCACACTTCCTTCTAC[G/A]GTGAATCTCAATCCTGACAC	TCTCCATATTCACACTTCCTTCTA	GTGTGTCAGGATTGAGATTCAC	19504	9
WIAF-2785	G/A	A	G	ATGGTCCACTGGGTGCCATG[G/A]CTGCAGTAACTGGTTCAGGG	TCCACTGGGTGCCATG	CCCTGAACCAGTTACTGCAG	19281	12
WIAF-2787	A/G	A	G	ACTGGTTCAGGGATGGTCCT[A/G]TGACCTTAGATTATCCAATC	TGGTTCAGGGATGGTCCT	CCGATTGGATAATCTAAGGTCA	19505	12
WIAF-2788	T/C	C	T	TTTGGGCTCTATACTATAAA[T/C]CACAGACTAGAGGCTCAAGA	CCAGTTGTTTGGGCTCTATACT	AACTCTTGAGCCTCTAGTCTGTG	19240	4
WIAF-2789	C/A	A	C	TTATATGTCACAAAACTTTT[C/A]AATGATTCTACTGAAGAAGT	GAATGTGTTTATATGTCACAAAACTTTT	TCTGAAACTTCTTCAGTAGAATCATT	19506	14
WIAF-2791	C/T	C	T	TTAAGACAGTAGCAGTGGGG[C/T]AAATAGATGGAGGGCTGATT	TTTAAGACAGTAGCAGTGGGG	CAAATCAGCCCTCCATCTATTT	19185	4
WIAF-2802	G/A	A	G	AGTGCTCCTGTTTTCTAGAG[G/A]TGCCTCTATTTTCTCTTTAC	GATTTTAGTGCTCCTGTTTTCTAGA	AGTGAGTAAAGAGAAAATAGAGGCA	19507	15
WIAF-2803	G/C	C	G	AAAGGAGTCCACTGGAGAGA[G/C]CCTGAGGCCATGGAGAGAGA	CTAAAGGAGTCCACTGGAGAGA	CTCTCCATGGCCTCAGG	19252	3
WIAF-2804	A/G	A	G	TTCTTTCTAAAAAATAAGCA[A/G]CTTAAATTGACCCAGGAAAG	TGAGACCCAATATTTTCTTTCTAAAA	GCTTTCCTGGGTCAATTTAAG	19283	3
WIAF-2805	C/T	C	T	AGCTTCTCTCTGCCTTGGCA[C/T]TTACGTCTCAGGCTCTGAAC	CTTCTCTCTGCCTTGGCA	TGTTCAGAGCCTGAGACGTAA	19508	10
WIAF-2809	G/A	A	G	AGAAAAGCTTGTGGATGCTG[G/A]GCACCATCCTATAGTGACTA	AGAAAAGCTTGTGGATGCTG	TCATTAGTCACTATAGGATGGTGC	19509	8
WIAF-2810	G/A	A	G	TCTTAGGTTACAGTAAATGA[G/A]AGTTTGTGCCAGAAGTGAGA	GGTCTTCCATCCTTTCTTAGGT	CACCCTTTCTCTAGTTTTCTCACT	19268	6
WIAF-2813	T/C	C	T	AAGGGCAGTGTAAGAAGAGA[T/C]TTTTTCCTTTAACAACATCT	GGAAGGGCAGTGTAAGAAGAGA	TGTCTTAGATGTTGTTAAAGGAAAAA	19325	18
WIAF-2815	G/T	G	T	GTAAACTTCTTATTTGGGTT[G/T]ATTTTGACTCTTACAGGTGT	AGTGATTGTAAACTTCTTATTTGGG	GAAAGACACCTGTAAGAGTCAAAAT	19510	7
WIAF-2816	A/G	A	G	GCAGGAAGTTCAGAGAGCAC[A/G]GCAGACCACTTGCAAATCAC	GCAGGAAGTTCAGAGAGCAC	TGATTTGCAAGTGGTCTGC	19152	2

WIAF 命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi 编号	NwLib chrom
WIAF-2817	C/T	C	T	ACTGCCTGGCTGTGGCTCGA\[C/T\]GCACGGTGATGGGGACCTCA	CCTGGCTGTGGCTCGA	TCCCCATCACCGTGC	19275	7
WIAF-2818	G/A	A	G	TAGATTCAGGAAATAAGGGT\[G/A\]AAAATGTAGGTGAAGTTAGG	GCCTAGATTCAGGAAATAAGGG	CATGACCTAACTTCACCTACATTTT	19511	1
WIAF-2820	C/A	A	C	TTTACAGCAAATTCTTTAGA\[C/A\]CAACTACCTAAAATGTTTTC	TCATATCTAATTTACAGCAAATTCTTTAG	CAGAAGAGAAAACATTTTAGGTAGTTG	19512	
WIAF-2822	C/A	A	C	CTCTCATGGGAGTGGAGAGC\[C/A\]ATTAGCATATAATCTAAAAT	TCTCATGGGAGTGGAGAGC	CCAGATTCATGATTTTAGATTATATGC	19513	11
WIAF-2823	G/C	C	G	GCAATTGCCCCAGGACCTAG\[G/C\]CTAGTCCAGCAGAGGCCACT	AATTGCCCCAGGACCTAG	TGGCCTCTGCTGGACTAG	19514	11
WIAF-2824	C/T	C	T	ATTAATACTGCCATGTGAAT\[C/T\]GTAAACACTTGGACTTGCCT	ATAAACATTAATACTGCCATGTGAAT	TGAGGCAAGTCCAAGTGTTT	19515	18
WIAF-2827	A/G	A	G	AGTTAAATGGTTGTTAAACT\[A/G\]ACAAACAAGAGAAGGGAAAC	TGATCAAGAGTTAAATGGTTGTTAAA	CGTTTCCCTTCTCTTGTTTG	19516	2
WIAF-2828	A/G	A	G	GAAAAGGTGAGGGCCTTTAT\[A/G\]TATCACCAGAGCCTTGAGCT	GAAAAGGTGAGGGCCTTTAT	AGCTCAAGGCTCTGGTGATA	19517	3
WIAF-283	C/T	C	T	TTGGCATTCAATAGCCTATC\[C/T\]AACTCCATGTGGGAGTTTCA	AGTTGGCATTCAATAGCCTATC	TGAAACTCCCACATGGAGTT	2624	8
WIAF-2832	G/A	A	G	GATTTATTGTAGAGATATGC\[G/A\]TTAAGAAGTCCACGTGTCCA	CATTAAGTCTGATTTATTGTAGAGATATGC	GTGGACACGTGGACTTCTTAA	19229	X
WIAF-2834	C/T	C	T	TGCTGGGGCCATCCCTGCCT\[C/T\]AACTCACATGCCCAGCAGCC	GGGCCATCCCTGCCT	GCTGCTGGGCATGTGAGT	19274	
WIAF-2837	G/C	C	G	GCCTCTGAGATTTTGAGATT\[G/C\]TGTTACATGCAGTTGATCTT	TGAAGCCTCTGAGATTTTGAGA	TTGTAAGATCAACTGCATGTAACA	19289	
WIAF-2839	T/A	A	T	ATTTGGCAGTGATTTAAAGA\[T/A\]AACAGCTAAAGGGCAAATCT	GATATAGATTTGGCAGTGATTTAAAGA	GAGATTTGCCCTTTAGCTGTT	19518	
WIAF-2841	T/C	C	T	AAGGAAGATTGGACAGATGA\[T/C\]AGGCAGAGGGGCAGGTAGGT	GGTTAAGTAAGGAAGATTGGACAG	CCTGCCCCTCTGCCT	19215	3
WIAF-2847	G/T	G	T	GGTGGGGCTCTTCTAAATGG\[G/T\]CTGGGATAATTCAGACATTA	GTGGGGCTCTTCTAAATGG	TCATTAATGTCTGAATTATCCCAG	19213	1
WIAF-2848	T/G	G	T	GACTTTTCTTTCTTCAAAAA\[T/G\]AGCATAGCTGAGCAATAAAG	GCAATGACTTTTCTTTCTTCAAAA	CCTCTTTATTGCTCAGCTATGC	19334	8
WIAF-2849	C/T	C	T	AGCTATATGACTGTAGGCAA\[C/T\]ATGGAGGAATATTACTTGGG	AGAAAAGCTATATGACTGTAGGCAA	TTTCCCAAGTAATATTCCTCCA	19314	2
WIAF-2853	C/T	C	T	CTTGTCCCTCCCCCTTCCTG\[C/T\]GCTGGAGCAGCTCTCTCTGG	CCCTCCCCCTTCCTG	CAGAGAGAGCTGCTCCAGC	19260	2
WIAF-2857	G/C	C	G	AAGCTTTGGAAAAGCCATCT\[G/C\]CTCAATTCCCAGTGGCTTAT	AAGCTTTGGAAAAGCCATCT	CATAAGCCACTGGGAATTGAG	19519	19
WIAF-2858	C/A	A	C	GCTTACCCTTGAACAGGGAA\[C/A\]CTTCACTGTCCCAGCTGGGC	GCTTACCCTTGAACAGGGAA	CCAGCTGGGACAGTGAAG	19253	17
WIAF-2859	T/C	C	T	TGCATGACACAATCATAGTC\[T/C\]GTTCTCATATTCTTTCTTTT	CAGTGCATGACACAATCATAGTC	CCCCTCATAAAAGAAAGAATATGA	19155	5
WIAF-2860	G/A	A	G	TGCCTTCTACCCCAGGCTGT\[G/A\]CAATGACGAGTACTCTTGGA	CCTTCTACCCCAGGCTGT	TGTCCAAGAGTACTCGTCATTG	19520	16
WIAF-2861	C/T	C	T	CGAGGACATCATGGTCCAGC\[C/T\]CGAGGGAGATTTATCTTTGA	GAGGACATCATGGTCCAGC	CAATCAAAGATAAATCTCCCTCG	19343	16
WIAF-2866	A/C	A	C	CAGACATGGATTTTGAATTC\[A/C\]GTTTTTCTTGTTCCAGTTTT	AATTACAGACATGGATTTTGAATTC	CCAAAACTGGAACAAGAAAAAC	19203	1
WIAF-2867	C/T	C	T	ATAAGGAAATTAGCGTCCTT\[C/T\]TGGACTTCTAGAGGATACTG	AAAGATAAGGAAATTAGCGTCCTT	TTCCAGTATCCTCTAGAAGTCCA	19162	14
WIAF-2868	A/G	A	G	TGAAACAGTATTAAACAGGA\[A/G\]ATTCAAGGCAGAGTTAGCCC	TTCCTTTGAAATAAATTGAAACAGT	GGGCTAACTCTGCCTTGAAT	19521	14
WIAF-2869	C/G	C	G	CATGGATTGTTAACTGCTTT\[C/G\]CGATGTTGACAGCACTGCTG	AGCATGGATTGTTAACTGCTTT	GCAGTGCTGTCAACATCG	19522	2
WIAF-287	T/G	G	T	CATGAAGAGTGGGCAGTTCA\[T/G\]GTTTTATTAGTATATAAAAT	CATGAAGAGTGGGCAGTTCA	TCCTGTAAAGCCAATTTTATATACTAATAA	2627	5
WIAF-2871	G/A	A	G	AAGGTATTTTCTAAGCTTTT\[G/A\]GATTGAAGGACAATCAAGCA	TCTATAAGAAAAGGTATTTTCTAAGCTTTT	CTATGCTTGATTGTCCTTCAATC	19523	3
WIAF-2874	G/A	A	G	AAGAATAATTATAACATTTT\[G/A\]AGCACATATTAAGCACCAGG	AAGCAAGATCACACAGAGAAAAA	GCCTGGTGCTTAATATGTGC	19524	17
WIAF-2876	G/C	C	G	TGGGAGTCATTAACAGTTAA\[G/C\]AACTTTTTCAGTGGTAAAAA	CACCTGGGAGTCATTAACAGTT	GCTCTTTTTACCACTGAAAAAGTT	19525	11
WIAF-2877	A/G	A	G	AGAGTCAACAGGGCTGCCAA\[A/G\]GTCCCAAAGGGTCTGGAATT	TCAACAGGGCTGCCAA	TTCCAGACCCTTTGGGAC	19225	6
WIAF-2878	C/T	C	T	TATCAGATTATTCTTTATTG\[C/T\]TTCCAGAATACTGGGAGGCT	CACCTTTGTTTGTATATCAGATTATTCTT	GCCTCCCAGTATTCTGGAA	19236	11
WIAF-2887	C/T	C	T	ACAAAAAAAGTATTTTACTC\[C/T\]ATACAAATTTAAAAAATAAA	AAAACTCAGTTGTGCAACAAAA	TCATCCCCATTGTTTTATTTTT	19526	15
WIAF-2889	C/T	C	T	TCCAACAAAGGCTAAGTGGA\[C/T\]TCTCAGCAAGACTATATTAA	CACCTCCAACAAAGGCTAAG	TGGATACTTTTAATATAGTCTTGCTGA	19328	15
WIAF-2891	A/C	A	C	ATGAAGGCTTCGGGCTTCGA\[A/C\]TTCTCTCAAGTCCTAGCCTT	GGCTTCGGGCTTCGA	ACAAGGCTAGGACTTGAGAGAA	19527	3
WIAF-2892	A/G	A	G	GAAACACAGTTTATTACAGT\[A/G\]GGATTAGAGTACACAAACTC	CAACTTCAGAAACACAGTTTATTACAG	AGAATATAGGAGTTTGTGTACTCTAATCC	19528	6
WIAF-2894	A/G	A	G	CTTATCCTGCTGCTAGCTGA\[A/G\]TTACAATGCAGGCTGAACTC	TCTTATCCTGCTGCTAGCTGA	GGAGTTCAGCCTGCATTGTA	19529	11
WIAF-2897	T/A	A	T	TCTTCAGTACCCTCATTTGA\[T/A\]TTCTTAGTTCTTGCCCTTTG	CCTCTTCAGTACCCTCATTTGA	GCAAAGGGCAAGAACTAAGA	19157	11
WIAF-2907	T/C	C	T	TTTTTAATGAGTACCATTTA\[T/C\]GAGAGACACGAACAATTCTC	AACAAATTGAGAACTTTTTAATGAGTAC	GGTGAGAATTGTTCGTGTCTCTC	19530	4
WIAF-291	C/T	C	T	CAAACTGAACATTTAAATGT\[C/T\]ATCCATGTGAGGGCTCTAGA	GCTATTAGGCAAACTGAACATTTAAA	TCTAGAGCCCTCACATGGAT	2631	3
WIAF-2911	A/C	A	C	GGGCCTAGGCACTTAATCAG\[A/C\]GAACCTAAATTTAGATCTCA	GGGCCTAGGCACTTAATCAG	GGCAACAGTGAGATCTAAATTTAGG	19277	10
WIAF-2912	A/G	A	G	AAAGCCCCCTTCCTTGCCAT\[A/G\]GGGCTGGAACCTAGGCCTGG	CCCCCTTCCTTGCCAT	GGCCTAGGTTCCAGCCC	19278	1
WIAF-2914	C/T	C	T	GCAAGTGGATATTTTGATGA\[C/T\]GGTGATCTCCAAGCTATTGT	GGCAAGTGGATATTTTGATGA	GAACAATAGCTTGGAGATCACC	19531	8
WIAF-2919	A/G	A	G	GGTGTTGCCCAAGCAAATTC\[A/G\]TGGAGTAAGGGAGAAAAACA	TGTTGCCCAAGCAAATTC	GATGTTTTTCTCCCTTACTCCA	19532	14
WIAF-2920	G/C	C	G	TCTTCTGCTTTAATTTATTT\[G/C\]TGCTTAGCTATATTAACAAT	GAACAGTTTCTTCTGCTTTAATTTATTT	ACACAAATCTTATTGTTAATATAGCTAAGC	19533	4
WIAF-2921	G/T	G	T	ACTTACCTGCATTGCTTTTT\[G/T\]AAAATCTTCATATCTCTGCT	CATTACTTACCTGCATTGCTTTTT	CCCAGCAGAGATATGAAGATTTT	19315	1
WIAF-2922	C/T	C	T	TCCACTTCTTCTTCTACAAA\[C/T\]AGGCATGATCATTCCTTGTC	GGACTCCACTTCTTCTTCTACAAA	GGTAGACAAGGAATGATCATGC	19534	3
WIAF-2924	C/T	C	T	TTATTGCTGCTAATTCTGTC\[C/T\]GAACCACACGATTTTCCTCT	CTTAAATTATTGCTGCTAATTCTGTC	CAGAGGAAAATCGTGTGGTTC	19535	7
WIAF-2925	G/A	A	G	AGAAAACTGTGCTCAGTCCC\[G/A\]GGAAGGAGGGCAGCCTGATG	GAGAAAACTGTGCTCAGTCCC	GGCTGCCCTCCTTCC	19536	8
WIAF-2927	A/G	A	G	ATTCCTTCCTCTCTATCCAC\[A/G\]TTTTCAATCACATAGCAAAA	TCAATTCCTTCCTCTCTATCCA	TGGTTTTGCTATGTGATTGAAAA	19537	8
WIAF-2930	C/G	C	G	GGTGCAAGGCCCAGTGAGGT\[C/G\]TAGGTCCAGGGTCTCTCCAC	GCAAGGCCCAGTGAGGT	TGGAGAGACCCTGGACCTA	19199	2

续表

WIAF 命名	等位基因	等位基因 A	B	芯片中的位置	上游引物	下游引物	ncbi 编号	NwLib chrom
WIAF-2931	G/A	A	G	TCTCATGAGGACAGCACTCT[G/A]CAGTGACATCCTATTTTGTT	TTCTCATGAGGACAGCACTCT	GCAACAAAATAGGATGTCACTG	19248	15
WIAF-2935	C/A	A	C	TTTTTTTTCCTAAGAGTAAA[C/A]AACAAAAATTATCCCTGGGA	GGATGGAAGAATGTACACCTTAA	TTCCCAGGGATAATTTTTGTT	19538	5
WIAF-2936	C/G	C	G	TAACCAATACCATCATTACA[C/G]TCAAGCCATAGAGATCTAAT	TGTTTTAACCAATACCATCATTACA	AGATTCACATTTTCATTAGATCTCTATG	19255	4
WIAF-2937	A/G	A	G	GTTTTAAATGAGGGGAGAGG[A/G]CAATAACAGCACCATTTTCA	CAGTTTTAAATGAGGGGAGAGG	TGTGAAAATGGTGCTGTTATTG	19539	19
WIAF-2938	G/C	C	G	TCATTAATCAAACCCGACCA[G/C]AGTATAGCCAATATACAAAA	TCATTAATCAAACCCGACCA	CTCTCTTTGATTTTGTATATTGGCT	19302	12
WIAF-294	T/C	C	T	TTGGTTATTTGACTCCAAAA[T/C]TGAATAAGTATTGGGGAAGA	GCTTGGTTATTTGACTCCAAAA	GGATTCTTCCCCAATACTTATTCA	2634	14
WIAF-2941	T/C	C	T	TGAGCCACCCAAGGCGATCA[T/C]AAAAGCCAAAGTTACAGATG	CCAGGAAAAATGAGCCACC	TTCTACATCTGTAACTTTGGCTTTT	19303	5
WIAF-2943	A/G	A	G	ATAAACTCAATGATTTCACA[A/G]TATCTCAGCACTTTCCTGCA	CCTACTTTATAAACTCAATGATTTCACA	TGCAGGAAAGTGCTGAGATA	19540	5
WIAF-2945	T/C	C	T	TATTGAATATAAGTGCCATG[T/C]GACGTGGGAGCCTTCATTAA	AAAATTTATTGAATATAAGTGCCATG	TTAATGAAGGCTCCCACGT	19297	9
WIAF-2948	G/A	A	G	CCTTCCTGACTCAAGCCCAT[G/A]CTGAGTGCTTCTCCTATGTA	CTTCCTGACTCAAGCCCAT	CAGAGTACATAGGAGAAGCACTCA	19541	5
WIAF-2949	T/C	C	T	ACTGAGTGTCCGATTTGTGC[T/C]TATGAAACTCTTGATAGGAA	CTGAGTGTCCGATTTGTGC	TTAAACCTTCCTATCAAGAGTTTCA	19542	15
WIAF-295	G/T	G	T	TTGGTGCCTGCACACCCACA[G/T]TGGCAACTAAGTGTAATCTC	TGCCTGCACACCCACA	CCATATTTAGAGATTACACTTAGTTGCC	2635	14
WIAF-2950	G/T	G	T	CTTCTGCTCAGTGACTGATG[G/T]TCTATTTTCAGTTCCATGAA	CACTTCTGCTCAGTGACTGATG	AACCTTCATGGAACTGAAAATAGA	19543	9
WIAF-2952	C/T	C	T	TCTGCCATGAACTTAGGAAT[C/T]TAATTAGCTCAAAGGTCAGA	GGTCTGCCATGAACTTAGGAA	TGACTCTGACCTTTGAGCTAATT	19544	3
WIAF-2953	A/G	A	G	GACTAGTTTTTCTTAAAGGC[A/G]GTGCATTCACTTTGTAATGA	GCAAGACTAGTTTTTCTTAAAGGC	TTATTTCATTACAAAGTGAATGCAC	19257	9
WIAF-2956	C/T	C	T	GGGAAACTTAATCTGTATGA[C/T]GGAAAAGAGATAAGAAATAA	TGAATTGGGAAACTTAATCTGTATG	CCACATGAGAATTGTTTTTATTTCTT	19178	6
WIAF-2957	T/C	C	T	GTAGGAGAACTTTTTAGCAA[T/C]GGCCATCATGAGCTGTGTAA	GAAAGAGTAGGAGAACTTTTTAGCAA	ACACAGCTCATGATGGCC	19545	10
WIAF-2963	T/C	C	T	GTGATCCCTCTTCTGAGATG[T/C]CTTCTCTCAGACCCTACGAG	GAGTGATCCCTCTTCTGAGATG	CTCTCGTAGGGTCTGAGAGAAG	19546	20
WIAF-2964	G/A	A	G	GAGGACACGTGGGTGGGAAC[G/A]GGTGAAGATTCTCTCTGCTC	GACACGTGGGTGGGAAC	TGAGCAGAGAGAATCTTCACC	19547	4
WIAF-2967	A/G	A	G	GAACAGTGCTGGAAATCCCC[A/G]AGACAGCTGGATAGACGCCC	CAGTGCTGGAAATCCCC	GGCGTCTATCCAGCTGTCT	19287	16
WIAF-2970	A/G	A	G	CACCTAAGTCTGGCAGTCCC[A/G]GTATCAGATGCATGGGAGAA	CACCTAAGTCTGGCAGTCCC	CTTCTCCCATGCATCTGATAC	19182	5
WIAF-2971	A/G	A	G	AGCGTCATTGCTTTCCAAGT[A/G]GCCTAAGATGAGAACAAACT	GCGTCATTGCTTTCCAAGT	GGAATAGTTTGTTCTCATCTTAGGC	19160	18
WIAF-2975	A/G	A	G	TGTGGCTTAGGAAAACTCAT[A/G]GCAGAGCAGACAGTGAGGCC	GGTGTGGCTTAGGAAAACTCA	GCCTCACTGTCTGCTCTGC	19548	4
WIAF-2977	T/C	C	T	CTTGAAAGAGTTTCTGACAC[T/C]ATATTTTAATTTCTATATGA	TTCCTTGAAAGAGTTTCTGACAC	TCTTAGCATCATCAGAGGTCATAT	19284	6
WIAF-2978	G/A	A	G	GAATGAAGAGTCAGGTCAGC[G/A]GGCAGGGATTGGGGGGGTGG	TGAATGAAGAGTCAGGTCAGC	CCCCCAATCCCTGCC	19218	4
WIAF-298	G/A	A	G	TTTATTCTTCAGGACACTCA[G/A]TTCACATGCCACTCTCGTGA	TCTTTTTATTCTTCAGGACACTCA	CACGAGAGTGGCATGTGAA	2123	1
WIAF-2980	A/T	A	T	TCTCTTCTCGCATAATCCCA[A/T]CTGAAGCTTCCACTTGTTTT	TCTCTTCTCGCATAATCCCA	GAAAAACAAGTGGAAGCTTCAG	19321	8
WIAF-2983	T/C	C	T	GTACTGTGCTCCTCCGTGTA[T/C]CTCCCCGTGAGGCTGGGGTG	TGTACTGTGCTCCTCCGTGT	CCAGCCTCACGGGGA	19338	11
WIAF-2984	A/G	A	G	CTACATCTGTTTAGTGGATC[A/G]GTGAATAATAGAATGTGATT	GGAATTCTACATCTGTTTAGTGGATC	CTTTCAAAATCAATCACATTCTATTATT	19180	2
WIAF-2987	G/C	C	G	TGGATTTAGCTGGAGTTGTG[G/C]GCCAGCTCTTCTACAGAATT	GTGGATTTAGCTGGAGTTGTG	TCAATTCTGTAGAAGAGCTGGC	19164	6
WIAF-2989	C/T	C	T	CGCAGAGGGAATTTTATGTT[C/T]AGAACAGGAAGACTGACCTG	CCGCAGAGGGAATTTTATGT	GCTAGTCAGGTCAGTCTTCCTG	19549	
WIAF-2997	G/A	A	G	GGAAGGAAGGAGGCTTGCTT[G/A]GGGTCTGATCTGTAAAGCAC	GAAGGAAGGAGGCTTGCTT	TGTGCTTTACAGATCAGACCC	19267	16
WIAF-2999	A/T	A	T	ACACATGTGCTCTATCCACA[A/T]CAGCAGAGGTATTTTGTGTA	GGAAACACACATGTGCTCTATC	CACTTTACACAAAATACCTCTGCTG	19179	5
WIAF-3000	T/C	C	T	TCCTGGGTAAGAGGAGACAA[T/C]GATATTTCCATGGTCTTCAG	TCCTGGGTAAGAGGAGACAA	GCTCTGAAGACCATGGAAATATC	19226	5
WIAF-3001	A/G	A	G	TAATTTTCTTCTGAAATCAC[A/G]GAGCTCCAGTAACTGCGGTC	TGAGTTCTAATTTTCTTCTGAAATCA	CCGCAGTTACTGGAGCTC	19550	2
WIAF-3005	T/G	G	T	GTAATGCCATTTTTAAACAT[T/G]TTTCCATCCAGGTGTTCCAA	GCAGTAATGCCATTTTTAAACA	TGGAACACCTGGATGGAAA	19332	2
WIAF-3006	T/C	C	T	GATTCTTTTGCTGTGGGATC[T/C]GAGAAAAATTGCACATGAGA	GAGATTCTTTTGCTGTGGGA	AATTCTCATGTGCAATTTTTCTC	19327	9
WIAF-3007	T/C	C	T	CCTTTAGTTTTGCCTACAAC[T/C]TCACCATTATTTGGTCTATC	CAGCCTTTAGTTTTGCCTACAA	GACTTGATAGACCAAATAATGGTGA	19447	11
WIAF-3008	G/A	A	G	ATGTCTCACTTCCATCTCAC[G/A]GCTGGGATTGTCCCTCATTC	TGAACATGTCTCACTTCCATCT	TGAGGGACAATCCCAGC	19210	3
WIAF-301	A/T	A	T	GCATGTCCTAGACCCCATTG[A/T]CTTACGCAAACTCAATCAGC	GCATGTCCTAGACCCCATTG	GGCTGATTGAGTTTGCGTAA	2126	2
WIAF-3012	C/T	C	T	TCTGAGAAAGCAAGAAGAAG[C/T]TGATGCCATCCCCTTTACTC	GTGTCTGAGAAAGCAAGAAGAAG	AGTAAAGGGGATGGCATCA	19326	11
WIAF-3013	C/T	C	T	AGACTTGACTTGCTTTGTTT[C/T]TTTCCATTCTGTCTCCCAAA	CTAGATCTTAGACTTGACTTGCTTTG	TTTGGGAGACAGAATGGAAA	19551	18
WIAF-3015	C/T	C	T	ACCATTGGTTAGCTCTGACC[C/T]GATGCATGGACCTCCCTGGG	CACCATTGGTTAGCTCTGACC	CAGGGAGGTCCATGCATC	19214	18
WIAF-3016	T/C	C	T	TAGATTTTCTTTTCTTGCCA[T/C]GAAAAATATATTCTCTAGCT	TCATAGATTTTCTTTTCTTGCCA	GGCCAAGAGCTAGAGAATATATTTTT	19329	8
WIAF-3019	G/C	C	G	TTCACATAGGAATGGGCTAA[G/C]TATACTCTATCCTGGGGCTG	TGTTCACATAGGAATGGGCTAA	CCAGCCCCAGGATAGAGTAT	19341	11
WIAF-302	C/T	C	T	TGGCTTATATTCTATTTCTA[C/T]TTGACAGCACAGTTCTTCAC	GTTTTTAATGTGGTATTAGAAAAGTTTAAA	CAATTTTCAGATTGTCTATAGCAAAC	19335	4
WIAF-3022	G/T	G	T	CCTGAAGATGATGAGTCCTC[G/T]ATCACCTTCATCTTTGAACC	TTCCTGAAGATGATGAGTCCTC	GGGTTCAAAGATGAAGGTGA	19552	15
WIAF-3023	C/T	C	T	GCCTGGGCATTGAGTCCTGT[C/T]GTCTAGGCTTTGTGTACAGT	CTGGGCATTGAGTCCTGT	CAAGCACTGTACACAAAGCCT	19175	11
WIAF-3024	G/T	G	T	TAAAGAATGAGGTTATAGAA[G/T]TAGGAAGGAGCTGATCAGGA	CTGGACTAAAGTATAATGAAAAAGCTAA	TCCTGATCAGCTCCTTCCTA	19553	19
WIAF-3025	A/C	A	C	GGAGTTGGCCCCACTCCAAG[A/C]GGCCCCCAGATCCTCCCCTC	TTGGCCCCACTCCAAG	GAGGATCTGGGGGCC	19168	9
WIAF-3029	C/T	C	T	CTGCAGCAGTCTTTAATAAA[C/T]CACATCTGGTCCTCTGGCCA	AACCTGCAGCAGTCTTTAATAAA	GCCAGAGGACCAGATGTG	19554	10
WIAF-303	G/A	A	G	CATACTATTTCCCAATTAGA[G/A]CCATGTCATTTTTCAGAAAA	ATGACCATACTATTTCCCAATTAGA	GCTTTTCTGAAAAATGACATGG	2127	10
WIAF-3030	A/G	A	G	AGAAATGAGGTGCTGGGCCC[A/G]GATCATCTTGGAAGGCCCTT	TGAGGTGCTGGGCCC	GGGCCTTCCAAGATGATC	19555	17
WIAF-3033	A/C	A	C	TGTGCTGGGTACTTAGGAAC[A/C]AACTGTGGGCTATTCACCCA	GATGTGCTGGGTACTTAGGAAC	GGGTGAATAGCCCACAGTT	19276	8
WIAF-3035	G/A	A	G	AGGGATTTCTGAGAAGCGGC[G/A]TGTGCCTGGTGTCCCTCTGT	GGATTTCTGAGAAGCGGC	GAGGGACACCAGGCACA	19224	9

续表

WIAF命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi编号	NwLib chrom
WIAF-3036	G/T	G	T	TAATAAAGTGACCCCGAGAT[G/T]TAGAGGGGCAGAAAAAGTAT	AGATAATAAAGTGACCCCGAGAT	TCATACTTTTTCTGCCCCTCT	19556	9
WIAF-3038	C/T	C	T	GTGTCTTACTAGGCTCACAG[C/T]AAGGCAAAGCTGAATGTGAA	GATCAGTGTCTTACTAGGCTCACA	TTCACATTCAGCTTTGCCTT	19163	1
WIAF-3041	A/G	A	G	CCACAGTGGTCACATGATCA[A/G]TCCTGGTCTTTATACCATTA	GGAGCCTTAGTCGCCCCAC	GGAAGTAATGGTATAAAGACCAGGA	19228	20
WIAF-3042	C/T	C	T	TGGTACTTGAGAGTATCATC[C/T]GAAATATGTAGAAGAGAAAT	AAAGATGATGGTACTTGAGAGTATCA	GGGCTTATTTCTCTTCTACATATTTC	19313	1
WIAF-3046	A/G	A	G	AGGACTTGGAAATCACCAAC[A/G]ATGGTCATAATCACGGTTAG	CAGGACTTGGAAATCACCAA	TGCTAACCGTGATTATGACCA	19204	2
WIAF-3047	T/G	G	T	GTTCTACATATCACAGGAAG[T/G]CACTGCCAGGTCTTAGGCAG	CAGATATTGTTCTACATATCACAGGAAG	TGCCTAAGACCTGGCAGTG	19238	4
WIAF-3049	T/G	G	T	AGGGAGAGAAGGGGTGGAAA[T/G]GAATTAAGTGTAGTCAAGGG	GGAGAGAAGGGGTGGAAA	CATTTCCCTTGACTACACTTAATTC	19244	12
WIAF-305	T/C	C	T	TGAACTGTCTAGCAGTATTA[T/C]GCTATTAGCTATGTTTACAA	GCATCATATGAACTGTCTAGCAGT	GGACAAATTGTAAACATAGCTAATAGC	2639	2
WIAF-3051	T/C	C	T	GAAGGTACTGGGGTCAATAC[T/C]GGAGATAGAATCAACCAAAC	GTGAAGGTACTGGGGTCAATAC	TGTACTTCAGTTTGGTTGATTCTATC	19557	17
WIAF-3053	T/C	C	T	CTTACCATCTAATTCAAAAA[T/C]GCACTTAAGTACATAGACCA	AAACACTCTTACCATCTAATTCAAAAA	TGTCACTGGTCTATGTACTTAAGTGC	19190	11
WIAF-3054	G/C	C	G	GAAAACTTACAAAGTATTTC[G/C]TTTATCACATGCCTATCTGT	CCGCTATGAAAACTTACAAAGTATTT	AAGATAACAGATAGGCATGTGATAAA	19211	8
WIAF-3055	T/C	C	T	GGGTTTGGGGTAGGGAAAGA[T/C]GAAAATATACACTAGATGTT	GGTTTGGGGTAGGGAAAGA	GATCCATGAAAGTTAATAATTCTCTTATT	19298	11
WIAF-3058	T/G	G	T	TGTTTTTGTGACTTAAGGAC[T/G]GTCATCAAGAGAACCTTAAC	TGGTATGTTTTTGTGACTTAAGGA	CTGGTTAAGGTTCTCTTGATGAC	19232	10
WIAF-3059	C/T	C	T	TCATGCAAAATTCGAGATTA[C/T]TTTAATGTCTGTGACTTAGT	CTTCATGCAAAATTCGAGATT	AAAATACACACTAAGTCACAGACATTAAA	19299	11
WIAF-3063	C/A	A	C	TCTGTTCCACTATTATTGTC[C/A]TTGGAAACAGGGCCTGAATT	ACTATCTGTCTGTTCCACTATTATTGTC	TCAGGCCCTGTTTCCAA	19558	12
WIAF-3064	C/T	C	T	ACAGTCCAGATTTCCATAAC[C/T]GCTTTCTCTTACTAGGCGAA	CAAAACAGTCCAGATTTCCATAA	TCTTCGCCTAGTAAGAGAAAGC	19559	X
WIAF-3065	A/G	A	G	TGAGTGTCCTTCACAAAGGC[A/G]CAGCTGTCTCCACAGCCCTG	TGAGTGTCCTTCACAAAGGC	GGGCTGTGGAGACAGCTG	19273	22
WIAF-3067	A/G	A	G	ATTTATTTAAGGTTCTATTT[A/G]CACTGGTCGGATATTACTGA	GACAACTATGTTTAGATTTATTTAAGGTTC	CCCCACACAGTGTATTCAGTAA	19560	22
WIAF-3068	G/A	A	G	CACCTGAACAGCTCCTGGTC[G/A]TTCTGCAGACACAACTTTGA	ACCTGAACAGCTCCTGGTC	TCTCAAAGTTGTGTCTGCAGAA	19340	16
WIAF-3070	G/A	A	G	CCATTAGGGAAGGGCGATGC[G/A]GGCAGAGAAGAAGGTAAAGG	AGGGAAGGGCGATGC	TCTGCCCTTTACCTTCTTCTCT	19170	13
WIAF-3071	A/G	A	G	TGCAAGCCCCTCTGTCTTCC[A/G]GTAGCTTGTGATGACGGGTC	CAAGCCCCTCTGTCTTCC	GACCCGTCATCACAAGCTAC	19239	11
WIAF-3072	G/A	A	G	CATCCCAGAATGCATCTTGG[G/A]AAAGGATATCTTATCAGGCT	TCCCAGAATGCATCTTGG	GTCAAGCCTGATAAGATATCCTTT	19561	1
WIAF-3073	A/T	A	T	GGCCCTGGGCAGCTCTAAGG[A/T]AGTTTTGCTGGAGGAGGAGA	CCCTGGGCAGCTCTAAGG	CTCCTCCTCCAGCAAAACT	19562	17
WIAF-3074	T/A	A	T	TTTCAGGGACTAGGTTTTTG[T/A]GGTGTTCCCTGTTTGTCAAC	CGTTTCAGGGACTAGGTTTTTG	GTTGACAAACAGGGAACACC	19312	17
WIAF-3078	A/C	A	C	AAGACCTGAGCCAATAAGGA[A/C]CACCCCTCTTTGTGTTTAAT	AAAGACCTGAGCCAATAAGGA	CAATTAAACACAAAGAGGGGTG	19322	14
WIAF-308	G/C	C	G	CCTTAAAACAACTATCAACA[G/C]CTGCAACACAAACCACAGGC	AACAATCCTTAAAACAACTATCAACA	CCTGTGGTTTGTGTTGCAG	2641	10
WIAF-3081	G/T	G	T	GGGGGAGGATTCATCTCCAG[G/T]TTTTAAAAATAACACGCTTA	GGGGAGGATTCATCTCCAG	AGATAGTGTAAGCGTGTTATTTTTAAAA	19563	6
WIAF-3084	T/G	G	T	TTTCTCTTACAAAGGTGACT[T/G]TGGATATCCCCTTCTAGCCA	CATGTTTTCTCTTACAAAGGTGAC	TGGCTAGAAGGGGATATCCA	19337	13
WIAF-3086	C/T	C	T	TATCTTGCCCACAGAATGTC[C/T]ACCTGAAGGTGGGTAGTTGA	CTATCTTGCCCACAGAATGTC	TCAACTACCCACCTTCAGGT	19301	4
WIAF-3093	T/C	C	T	GGAATATGAAGAAATCAAGT[T/C]GCCTCTTCTGAAGGAGAGTC	AAAAGGGAATATGAAGAAATCAAG	TGACTCTCCTTCAGAAGAGGC	19448	9
WIAF-3095	G/A	A	G	TGCTCAAATTTCATCATTTC[G/A]GAGACATCTCTCCTAATCAC	TCTTTGCTCAAATTTCATCATTT	TATATGGTGATTAGGAGAGATGTCTC	19316	1
WIAF-3096	T/A	A	T	TTCGCACTGCCTGCAGGTAC[T/A]GCTGCCACTCACCTCTGAAT	CGCACTGCCTGCAGGTAC	TCAGAGGTGAGTGGCAGC	19285	13
WIAF-3097	G/A	A	G	AAGTTGGTGGTGATTTAGCT[G/A]CAAGATCAGTTTCTCAAATG	TTTAAGTTGGTGGTGATTTAGCT	TCTTATCATTTGAGAAACTGATCTTG	19311	6
WIAF-3098	G/A	A	G	CCACCAGGTCCCTTCCACCC[G/A]CTGTGGCCTTGGCGTCCTTC	CAGGTCCCTTCCACCC	ACGCCAAGGCCACAG	19205	9
WIAF-310	C/T	C	T	TGTGGGCAGGGACTGTGTCT[C/T]GTTCCCTGTTGGGTCCCCGG	TGGGCAGGGACTGTGTC	GGGACCCAACAGGGAAC	2128	10
WIAF-3103	A/G	A	G	ATCAGACAGAGCAGGAAATA[A/G]GGAATAAGGAGGATATGTCA	TGATATCAGACAGAGCAGGAAA	TGAGTGACATATCCTCCTTATTCC	19233	3
WIAF-3106	C/G	C	G	CACCACTTCAAATCACAAGA[C/G]TAGCTAACATATACTGAATG	CCACCACTTCAAATCACAAGA	TGTGATAAGCATTCAGTATATGTTAGC	19220	15
WIAF-3110	T/A	A	T	CCACCACGCAAATCATTCTA[T/A]TAACAATTCAAGCCAGGTGT	CCACCACGCAAATCATTCTA	GACACCTGGCTTGAATTGTT	19191	8
WIAF-3111	C/T	C	T	CAAGCTATCATTTGACTTTC[C/T]AGGGACTCTTCTAAGAATAA	GGATTATTAGTTCCTTAAATCCTTCC	AGTTCAAGGATAGTTTTATTCTTAGAAGA	19564	15
WIAF-3112	T/C	C	T	CAGGACAGGTCCTGAGGGTC[T/C]TAGAGCTGAGCCTCTTTTAA	GGACAGGTCCTGAGGGTC	CCTCTTAAAAGAGGCTCAGCTC	19234	3
WIAF-3114	G/A	A	G	CCTTGAAGAGCTCAATGAAT[G/A]CTGCCCACTAAGTGCCCTCA	GGAATCCTTGAAGAGCTCAA	GGGCACTTAGTGGGCAG	19221	15
WIAF-3118	C/G	C	G	AAGAGGTGGTAAGAAGTGTT[C/G]AGATTCTGCATATTTTGTTA	GCAGAAGAGGTGGTAAGAAGTG	CATTCTATTAACAAAATATGCAGAATCT	19565	4
WIAF-3119	G/A	A	G	ACTGAAAAAAAAAAAAAAAG[G/A]GCATGTAGAAAAATCCTATA	CTGTTTTCAAGTCCAGACTGAA	AGATCTGGTAATTATAGGATTTTTCTACA	19566	12
WIAF-3120	T/A	A	T	GGAGGCATGACATACTATTA[T/A]TAAGTTGATCTGTGTATTTG	TGTTGGAGGCATGACATACTATT	AGTTTCCAAATACACAGATCAACTT	19286	18
WIAF-3122	G/A	A	G	CCTGGGTGTGTGTTCACTGC[G/A]TCCTCTGCTGCCCTGTCTTG	TGGGTGTGTGTTCACTGC	GACAGGGCAGCAGAGGA	19342	1
WIAF-3123	G/T	G	T	GTTGATAAGTTGTTCATAAA[G/T]CCAAATAAACCATGCGCGAA	GGGGTTGATAAGTTGTTCATAAA	CGCGCATGGTTTATTTGG	19567	7
WIAF-3124	A/C	A	C	TATAAGAAAAGCCTGGAATT[A/C]GATACTCTAAAATGTCAACA	TGTGTATATTTATGTATAAGAAAAGCCTG	CCTTAAAATAACTGTTGACATTTTAGAG	19568	6
WIAF-3125	C/A	A	C	TTGGCCACTCCACTTAACCT[C/A]TCTGGGCTTCTTTGCTCATC	TGGCCACTCCACTTAACCT	TGAGCAAAGAAGCCCAGA	19269	11
WIAF-3126	A/C	A	C	CTTTTTCAAAAAATACCTAT[A/C]GGGCCGGGCGCAGTGGCTCA	GGATAATGGGCTTTTTCAAAA	CTCGTGATGCGCCCC	19192	15
WIAF-3129	G/T	G	T	GTAGCTACTATCTTTCTATA[G/T]ATTTACCTCCTGCCCCATC	AGTTTGAGTGGTAGCTACTATCTTTCTA	TGGGGCAGGAGGTAAAAT	19569	11
WIAF-313	C/T	C	T	ACTTTATGCCATGGTGCTTT[C/T]AGTTTAATGCTGTGTCTCTG	ACTTTATGCCATGGTGCTTT	TTTGAGAGTTTATCTGACAGAGACA	2644	
WIAF-3130	T/C	C	T	CAGCAGGAGACACTCAACTT[T/C]CACTTGTTGCCCCGTTCAGT	GCAGCAGGAGACACTCAACTT	GAACGGGGCAACAAGTG	19570	10

续表

WIAF 命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi 编号	NwLib chrom
WIAF-3133	T/A	A	T	GCAGTCATCTTTTCTCTGCA[T/A]AAATACCTATGCCTTTACAG	AGCAGTCATCTTTTCTCTGCA	CACAGTCTGTAAAGGCATAGGTATTT	19571	5
WIAF-3136	G/C	C	G	GCCCCGTTCAGTAAAAGCAA[G/C]ATCCTAAATGAAAAACTGAC	CCCCGTTCAGTAAAAGCAA	CGTTGGTCAGTTTTTCATTTAGG	19572	10
WIAF-3137	T/C	C	T	TAGTTAGGGGTCCATTTACC[T/C]GAAGAGGTATCTGGGATTGT	CCTAGTTAGGGGTCCATTTACC	AGAACAATCCCAGATACCTCTTC	19188	5
WIAF-3139	C/A	A	C	GTCTCCCTCCCACTCTCCCA[C/A]AGACATCACTAGTCAAAAGG	CCCTCCCACTCTCCCA	TTGCCTTTTGACTAGTGATGTC	19270	11
WIAF-314	C/T	C	T	TCTGCCCAATGCCTAGAATA[C/T]AGTGGGTCCCTAATAGTTAT	TCTGCCCAATGCCTAGAATA	GGAACTAATAACTATTAGGGACCCA	19262	9
WIAF-3140	T/G	G	T	AGTGGCAGCAGTTGTAGAAC[T/G]TGGTATATGCATGTGCTTCT	ACTAGTGGCAGCAGTTGTAGAAC	AGAGAAGCACATGCATATACCA	19261	12
WIAF-3143	T/C	C	T	TAAAGATGGGCAAAGAAATA[T/C]CAGGATAGTGCTGGTAATGG	GGGAAATAAAGATGGGCAAAG	TCCATTACCAGCACTATCCTG	19573	2
WIAF-3144	A/G	A	G	AAAGAAAGAAACACATTTGT[A/G]TAAGTAAATGATTCCCAAAT	TCAGGGTAAAAGAAAGAAACACA	AGTTGATATTTGGGAATCATTTACTT	19258	4
WIAF-3145	C/G	C	G	TTCTCCTATTTAAAAAGCCA[C/G]ATAAATTTAAACATAAATAT	TGGTTCTCCTATTTAAAAAGCCA	GCTGTGTGGGATATTTATGTTTAA	19319	X
WIAF-3148	C/G	C	G	CAGCTAGTGTAGGCATTTTT[C/G]TGCTTTTCTAGTTGCTTTTA	ACACAGCTAGTGTAGGCATTTTT	GCAGATATAAAAGCAACTAGAAAAGC	19574	6
WIAF-315	A/G	A	G	ATATATCAAGGTAAAGTCCA[A/G]TACAAAAAAACAGCATTTCC	CCATATACATATATCAAGGTAAAGTCCA	GCCATAGGAAATGCTGTTTTT	2645	8
WIAF-3151	A/C	A	C	TAATGTTTAGTAGGCGTTAG[A/C]TCTCAGATTCTGAAGTTCAG	TTAACTGATAATGTTTAGTAGGCGTTAG	CCCTGAACTTCAGAATCTGAGA	19304	3
WIAF-3153	A/G	A	G	GTCCTTCCTCAGCAGCACTC[A/G]AGAGGAAACCCAGCAAGAAA	CCTTCCTCAGCAGCACTC	TTCTTGCTGGGTTTCCTCT	19208	9
WIAF-3156	T/C	C	T	ACCTATGTGGGGAGTTTGGC[T/C]GATACGGTTTTATGAGGCCC	CCTATGTGGGGAGTTTGGC	GGGCCTCATAAAACCGTATC	19247	8
WIAF-3157	C/T	C	T	ATGTAATCAAGTTCCTTTAC[C/T]GGGCTCTCGGAGGAGAAATA	GCCAAATGTAATCAAGTTCCTTT	TCTCCTCCGAGAGCCC	19575	8
WIAF-3158	G/T	G	T	AAACTAAGATAAGATTCTAG[G/T]AAATGATTGCTGCTACCCAC	TGCTAAAATTAATGTCTTGTAAACTAAGA	GTGGGTAGCAGCAATCATTT	19576	6
WIAF-316	C/T	C	T	GTCAGGGTAAAGCATTATGA[C/T]AGACACAAAGACAAAGAGGT	TGGTCAGGGTAAAGCATTATGA	GCAACTTTACCTCTTTGTCTTTG	2129	4
WIAF-3161	T/C	C	T	ACTCCAGGATACATGTCTAT[T/C]TTGAGCCAAACAGCAGTTTA	CTAACAACTCCAGGATACATGTCTAT	AAACTGCTGTTTGGCTCAA	19165	6
WIAF-3164	A/T	A	T	GAGCCTGTGGTTTTCTAAAA[A/T]CATGATGCTTTCTTGAATAA	TGAGCCTGTGGTTTTCTAAAA	CAGAATTATTCAAGAAAGCATCATG	19158	6
WIAF-3166	G/C	C	G	CTAATCACTGGGGAACAAAA[G/C]AGTAACATTTTAGGTACTGT	GACTAATCACTGGGGAACAAAA	TGAAAATTCACAGTACCTAAAATGTT	19577	6
WIAF-3168	C/A	A	C	GACAATACAGCAAAATTGTT[C/A]CCCTTTCACTGTTCAACAAA	TTACTTATGACAATACAGCAAAATTGT	GATGGGTGCATATTTGTTGAA	19578	2
WIAF-317	A/G	A	G	TAAAAAGATCTGCTTATCCT[A/G]TATATCCACATAACTCTAGT	GGACTTAAAAAGATCTGCTTATCCT	AGAAACTTGCTAAATATTTTATGTAACACT	2646	15
WIAF-3172	G/A	A	G	ATGAAGTCACAGAGTAGACA[G/A]GACAAGGGGATAAAACCGTA	CAGCATGAAGTCACAGAGTAGACA	CTACGGTTTTATCCCCTTGTC	19579	
WIAF-3176	A/G	A	G	GTCTGCTCATCTCGTTGAGT[A/G]TGTAGGAGGGATGCAGCTTC	TGTCTGCTCATCTCGTTGAGT	AAGCTGCATCCCTCCTACA	19580	3
WIAF-3178	C/T	C	T	ATGAACTTGTCAAGTTTGAT[C/T]ATGTGTTCCCTTACCTTTAA	TGCATGAACTTGTCAAGTTTGA	GAATTTTAAAGGTAAGGGAACACA	19330	18
WIAF-3180	G/A	A	G	ATTGAGGTAGAAGGAGTACA[G/A]CCCATTAGCAATCATTAACA	GGAGACATATTATTGAGGTAGAAGGA	GGTGTTAATGATTGCTAATGGG	19195	18
WIAF-3184	A/G	A	G	TGGACTTTTGTAGTTCATGA[A/G]TTTTGATTGTATTGTTACAA	TGATGTTCTTGGACTTTTGTAGTTC	TTGCTTTGTAACAATACAATCAAAA	19196	14
WIAF-3185	A/G	A	G	ACATTGTCCAATTATGAAAT[A/G]TGAGGTGAGGCCTGAAATAC	TGAAACATTGTCCAATTATGAAAT	GTATTTCAGGCCTCACCTCA	19581	14
WIAF-3186	T/C	C	T	GGACAAGAAACAGAAACAGA[T/C]GATAAGGCCACAGAAATGCC	ATGGGACAAGAAACAGAAACA	GGCATTTCTGTGGCCTTATC	19272	11
WIAF-3187	G/A	A	G	TACGAGGCAGGGCTAGACCC[G/A]GAAGCCTTGTCTGTGTGATT	GAGGCAGGGCTAGACCC	GAATCACACAGACAAGGCTTC	19582	7
WIAF-3188	G/T	G	T	GAGAAACAGGTTGCACGATA[G/T]GTGGAGAAACAACAGGTTGC	GGAGAAACAGGTTGCACGAT	GCAACCTGTTGTTTCTCCAC	19583	20
WIAF-3189	G/T	G	T	CTCACTGAGGTGTTTAAAAT[G/T]GAAACAGAATCAAGTAGCCA	GGCTCACTGAGGTGTTTAAAA	CCATGGCTACTTGATTCTGTTTC	19167	13
WIAF-3190	G/A	A	G	AAAAGGGTTTTGGAGACTGT[G/A]TCTAATGAAGGTAAATAAAT	AGAAAAGGGTTTTGGAGACTG	CCACATTCCAAGTTTATTTATTTACC	19254	6
WIAF-3191	G/T	G	T	GACAAAAGGAGTGTTTGGAA[G/T]GATGAAAGGTCATGAGAAGC	TGACAAAAGGAGTGTTTGGAA	GGCTTCTCATGACCTTTCATC	19584	6
WIAF-3194	T/C	C	T	TTGGTACATTTATCTTAATC[T/C]CTGCATCTTTGAACCAGGTT	CTGTTTACCACTTCTGCTTGG	AACCTGGTTCAAAGATGCAG	19339	7
WIAF-3197	A/G	A	G	GTCTCTGTATCCTGGGGAGT[A/G]GGATTTGCTGTTATCCTCAG	CAGTCTCTGTATCCTGGGGAG	GCTGAGGATAACAGCAAATCC	19235	17
WIAF-3198	G/C	C	G	AACTCCTAGTAATCTCTCAA[G/C]ATCCAGCTCAAATGTGGCTT	AGCCAACTCCTAGTAATCTCTCAA	GCCACATTTGAGCTGGAT	19585	5
WIAF-3199	T/G	G	T	CTCCTTTCCTCCCTTGCATG[T/G]ACCTTTGCTGTTGTCTCATC	CCTTTCCTCCCTTGCATG	TGATGAGACAACAGCAAAGG	19586	5
WIAF-3203	G/A	A	G	AATTAAGCCTCTCTCAATAT[G/A]GGGAAAGTCAGAACACTGGA	AGAACAAATTAAGCCTCTCTCAATA	TCCAGTGTTCTGACTTTCCC	19587	11
WIAF-3204	G/C	C	G	TGGCTGGATACAAAGGTCTC[G/C]AAGGCCTAGACTTCCACAGT	TGGCTGGATACAAAGGTCTC	CACTGTGGAAGTCTAGGCCTT	19588	18
WIAF-3206	T/C	C	T	AGTGATGGTTTCAGAGCCTG[T/C]CCTTGTATGGTGTGTTGAAG	AGTGATGGTTTCAGAGCCTG	TTCTTCAACACACCATACAAGG	19194	9
WIAF-3208	C/A	A	C	GATTGTTGAGTGAGGAGTGC[C/A]TAGATTTAAAATATATATTG	TAGATTGTTGAGTGAGGAGTGC	CAGTCCTTAGACTTCTCCAATATATATTT	19589	7
WIAF-3209	T/G	G	T	TAAAGAAAAAGGGTGATTAG[T/G]TGTATCAAATGTTGCTAATC	CAAATATTTAAAGAAAAAGGGTGATTAG	GCCTGATTAGCAACATTTGATACA	19309	7
WIAF-3210	A/G	A	G	AATGGCATCTGTATTCACAT[A/G]TTAGATATTAGACATTGTGT	TGAAATGGCATCTGTATTCACA	AAAATTGTAGCATCACACAATGTC	19171	5
WIAF-3212	A/T	A	T	AGAAACCAGGTGAGTAAGGG[A/T]CAGGCTCCAGAGACTTCTGG	CTAGAAACCAGGTGAGTAAGGG	CCAGAAGTCTCTGGAGCCT	19590	7
WIAF-3213	G/A	A	G	GGGGTCCCTGGCTGACTCAG[G/A]GAGGGCCGGCTCGGGTCTGC	GGTCCCTGGCTGACTCAG	CAGACCCGAGCCGGC	19209	10
WIAF-3214	T/C	C	T	GTGAGATTTCATCAGCCTGG[T/C]GGGTAAACATGGCTGTTTGC	GGTGAGATTTCATCAGCCTG	GCAAACAGCCATGTTTACCC	19181	1
WIAF-3215	C/T	C	T	TTCTTCCTTTTACTAGGGCT[C/T]TAGTCCCCATAGTTTCTGAG	CACTTCTTCCTTTTACTAGGGCT	AGAACTCAGAAACTATGGGGACT	19280	1
WIAF-3216	A/C	A	C	ACCCCGGTGTCTCATCCAGG[A/C]CTGGACACGTGGTGAGAGCT	CCGGTGTCTCATCCAGG	GCTCTCACCACGTGTCCAG	19200	2
WIAF-3218	G/A	A	G	AATGAGAAGGTAGGAACACA[G/A]TATAGCTTAACATAAGAGCA	TCTTAAATGAGAAGGTAGGAACACA	GTACAAGTCTGCTCTTATGTTAAGCTATA	19320	8
WIAF-3219	G/A	A	G	TTAAAAGGACTTGCAGGCTC[G/A]CTTAGCACTTACTAGTAAAG	TTAAAAGGACTTGCAGGCTC	GGAATTCATTACTTTACTAGTAAGTGCTAA	19202	6
WIAF-3220	A/G	A	G	TTGGGGTAAATGCTGTGGTT[A/G]AATATCAGTAGAGGAACTGG	TGGGGTAAATGCTGTGGTT	GCAACCAGTTCCTCTACTGATATT	19216	20
WIAF-3221	G/A	A	G	AGGAACTGGTTGCAGGGCTT[G/A]GAGGCTTTGCAGGATGTTAG	GAACTGGTTGCAGGGCTT	CTAACATCCTGCAAAGCCTC	19217	20
WIAF-3223	G/A	A	G	TTAAAAGGGGATAACCCGGC[G/A]TACTTGGTGATTGTCTGATC	AAGGGGATAACCCGGC	TTTGATCAGACAATCACCAAGT	19591	2
WIAF-3224	G/T	G	T	GAATGAATACAGGGAGCTAG[G/T]TCTATTTCGAAAAGAAAGGT	GAGAGAATGAATACAGGGAGCTAG	ACTCAACCTTTCTTTTCGAAATAG	19305	2
WIAF-3225	T/C	C	T	TGGACTCTTACATTCCTTCA[T/C]GTACACTGCTATTAAATGAG	AATCTGGACTCTTACATTCCTTCA	CATCGCTCATTTAATAGCAGTG	19592	4

续表

WIAF 命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi 编号	NwLib chrom
WIAF-3227	A/G	A	G	GGTTGCAACTTGGCAGCATC[A/G]CCTTCCAAACAAGAAACAAT	TGCAACTTGGCAGCATC	CAATTGTTTCTTGTTTGGAAGG	19593	14
WIAF-3228	G/C	C	G	GAGTAAGTTTGACACAGTTT[G/C]ACTTTAACATGTCAGTGAAA	AGGTTGAGTAAGTTTGACACAGTTT	CCTTCAACTTTCACTGACATGTT	19306	2
WIAF-3231	T/C	C	T	CTGTGTCCTATAGAGAGTCA[T/C]GGTCCCATCCCCTGGAACCA	TTGTTCTGTGTCCTATAGAGAGTCA	CCAGGGGATGGGACC	19594	8
WIAF-3232	G/A	A	G	ATCAGATTTTACGTGTAGCA[G/A]AGCCACAGATGTGATGATGG	AATTCAATCAGATTTTACGTGTAGC	CCATCATCACATCTGTGGCT	19333	1
WIAF-3233	C/G	C	G	TGAAGCAACATAAGCCAAAT[C/G]TTTGTGTTTATCTTACACGA	CTGAAGCAACATAAGCCAAAT	TCTCTTTCGTGTAAGATAAACACAAA	19153	1
WIAF-3235	C/A	A	C	GAACAAGCTTCTTACTACAA[C/A]CTGTACAGCACCAGCCTCAT	TGGAGAACAAGCTTCTTACTACAA	TGAGGCTGGTGCTGTACA	19595	3
WIAF-3236	G/C	C	G	AGCACATGTCTTAAAGCTTC[G/C]TCTCTGAGTAGGATCTGTGT	GCAGCACATGTCTTAAAGCTTC	CCCACACAGATCCTACTCAGAG	19596	11
WIAF-3237	C/T	C	T	ATTAATTATCTCTAAGCTTC[C/T]TTAAATTCCTAAGATTTTAA	GAGATTGGATTAATTATCTCTAAGCTTC	ACAGAGAATATTTTAAAATCTTAGGAATTT	19597	14
WIAF-3238	T/C	C	T	GTCATCGCTGTCACAGAGCT[T/C]GCAATCTGTCTGTCTCCGCT	TCATCGCTGTCACAGAGCT	TGATGTATAATAGCGGAGACAGAC	19174	7
WIAF-3241	C/A	A	C	TACAAAGAGAATTCAGACGA[C/A]ATCAGCAAGATGGTGGGATA	TCACAAAACTAATGTACAAAGAGAATTC	TCCCACCATCTTGCTGAT	19242	
WIAF-3245	G/A	A	G	TCTATCCTTCAAGCCTAGGG[G/A]TGGTGAAAATTTCATCACTA	CTCTATCCTTCAAGCCTAGGG	CCCTGTGTTAGTGATGAAATTTTC	19598	6
WIAF-3246	T/G	G	T	TAGACAAACTGAAATGCACT[T/G]AAATAAATGGATGAGGAAGT	CAATTTAGACAAACTGAAATGCA	AGAAACTTATACTTCCTCATCCATTT	19189	5
WIAF-3247	C/G	C	G	TGGTGGGCTGTGCTGCCTGA[C/G]GACAGCTGGGTATGACTCTC	GGGCTGTGCTGCCTGA	GAGAGAGTCATACCCAGCTGTC	19599	1
WIAF-3249	C/T	C	T	ATGATGCAAACTAGTGACCA[C/T]CACTTACTCCTTTCTGTACC	CATGATGCAAACTAGTGACCA	CGTGGTACAGAAAGGAGTAAGTG	19600	14
WIAF-3250	C/T	C	T	TCACCACCATCAGTAGACCA[C/T]GTCAAAAAAGAAGGGCTGTG	GACTCACCACCATCAGTAGACC	AAACCACAGCCCTTCTTTTT	1347	2
WIAF-3253	G/A	A	G	TTCCAAGCCAGACTGCCTGA[G/A]TATCACTTAAACCTAGAGTG	TTCCAAGCCAGACTGCC	TTTGGATGATAAATTGTACAGTCAC	1350	8
WIAF-3256	T/G	G	T	GGTTCCCGGAAGAAAAATTT[T/G]AAGACACATGGGTTAATGCA	GGTTCCCGGAAGAAAAATTT	CTGCATTAACCCATGTGTCTT	1353	4
WIAF-3257	T/C	C	T	GTTGTAAGACAGTTATTTTT[T/C]TGGGGACTGTTCTGTCTGGA	CCCTCGTTGTAAGACAGTTATTTTT	CCAGACAGAACAGTCCCCA	1354	4
WIAF-3260	A/T	A	T	TAGAACCAGAGAGGACAATG[A/T]TCTTGATTCCCGGGTTTCCC	CCATCCTAGAACCAGAGAGGAC	GAAACCCGGGAATCAAGA	1357	X
WIAF-3262	C/T	C	T	ATCTTATTAAAAGGCCCGTA[C/T]TGAAAGAAGAATCAGTAAAC	CAATATCTTATTAAAAGGCCCG	AAGAAGAGTTTACTGATTCTTCTTTCA	1359	5
WIAF-3263	G/A	A	G	TGTGCCAGATGCCTTCAAGG[G/A]TGACTAGATTGCCCATTTGA	GCCAGATGCCTTCAAGG	TTCAAATGGGCAATCTAGTCA	1360	1
WIAF-3265	G/C	C	G	CTGGATTGGCTTCGTTATTT[G/C]GTGGCCTGGCAACCTGTGGT	CCTGGATTGGCTTCGTTATT	CAGGTTGCCAGGCCAC	1362	4
WIAF-3268	A/T	A	T	ACACTCCTGCAGCTGAAAGG[A/T]GTCAATGCGAGCTGAGTTGC	CACTCCTGCAGCTGAAAGG	CAACTCAGCTCGCATTGAC	1365	13
WIAF-3270	C/T	C	T	ATCTCTCCTTACCTAGGGTA[C/T]GTTGCTTTCCAAACCAAGAG	CAGAAATCTCTCCTTACCTAGGG	CTCTTGGTTTGGAAAGCAAC	1367	16
WIAF-3271	T/C	C	T	GTATCAATGAATAGCTGATA[T/C]GCATTGTTGGGAGAATCATT	CCACTCTAAGTATCAATGAATAGCTGA	AATGATTCTCCCAACAATGC	1368	2
WIAF-3272	G/A	A	G	TGAGTTTGAGAGCAGAGTGC[G/A]TGGTATTTTTCTGGGCTTAA	CTGAGTTTGAGAGCAGAGTGC	CCTTAAGCCCAGAAAAATACCA	1369	1
WIAF-3276	G/A	A	G	AAAAACCCACACATGGGAGA[G/A]CCTCGGTAGGGAACCCAGCC	AAACCCACACATGGGAGA	TGGGTTCCCTACCGAGG	19201	19
WIAF-3277	A/G	A	G	AACTCTCTGTCCACTGGAGA[A/G]AAGTGCTTGCCACAGAAAAC	CCGAACTCTCTGTCCACTGG	TTTTCTGTGGCAAGCACTT	1373	6
WIAF-3279	A/C	A	C	CAAATAAAGTGGGCAGCCTG[A/C]GACTGGGCTATGACCTAGGG	AAATAAAGTGGGCAGCCTG	CCCTAGGTCATAGCCCAGTC	1375	8
WIAF-3280	C/T	C	T	CATCAGAAAATCCATACTGG[C/T]GAGAAACCTTATGAATGTAA	CAACATCAGAAAATCCATACTGG	ACATTCTTTACATTCATAAGGTTTCTC	1376	19
WIAF-3281	T/C	C	T	TGAGAGTTGTTATCCATGGA[T/C]CAGAAAAGAGCTCCCATGAA	AAGTGAGAGTTGTTATCCATGGA	TTCATGGGAGCTCTTTTCTG	1377	10
WIAF-3283	G/A	A	G	AGGTAGAGAAGCGGGTCTCC[G/A]AACCTAGGGTAATCTGGGCT	GGTAGAGAAGCGGGTCTCC	AAGCCCAGATTACCCTAGGTT	1379	14
WIAF-3285	C/T	C	T	AAAAAAGCTTCTCAGGTGAT[C/T]CTAAAGTATAACTAGATTTG	CAGGATTATGGGCATTTTTAAAA	TCTTCTGGTTCTCAAATCTAGTTATACTT	1381	7
WIAF-3287	C/T	C	T	CTTCTATGATGACTATGTAT[C/T]TATCCATGGGGAAGCTTAGT	GATTGCCTTCTATGATGACTATGTAT	AACTAAGCTTCCCCATGGAT	1382	15
WIAF-3289	A/G	A	G	CAATGAAAAGGGCTGAGGAA[A/G]TCCTCCAAAATGTGGAGACA	AATGAAAAGGGCTGAGGAA	TGTCTCCACATTTTGGAGGA	1384	2
WIAF-3291	C/A	A	C	TGTCCACATCATTCCATCTG[C/A]ATCGTCTTCCTACAAACAGT	TGTCCACATCATTCCATCTG	AAAAACTGTTTGTAGGAAGACGA	1386	11
WIAF-3292	T/C	C	T	AAGCACATTAAATTAATTCA[T/C]GAGTGAACATACAAGTATGA	GAAAATGTAAGCACATTAAATTAATTCA	GCATAGCTGTGATCATACTTGTATG	1387	7
WIAF-3293	T/C	C	T	AAGATATATTTTCCAAAACC[T/C]GAAAGAGGATAAATGGATGA	CGCAAGATATATTTTCCAAAACC	TGAGAAATCATCCATTTATCCTC	1388	1
WIAF-3294	A/G	A	G	TTTCTGAATATTTTAAAGAA[A/G]GACTTGCTCTCCTTATGCTC	GCTCCTTTTCTGAATATTTTAAAGAA	AGTTAGAGCATAAGGAGAGCAAGT	1389	10
WIAF-3296	T/C	C	T	TTTTGCCATTGCAGTGGACA[T/C]TTCTAAAAAGCAAATGCCTA	TGCCATTGCAGTGGACA	CCATAGGCATTTGCTTTTTAGAA	1391	12
WIAF-3297	T/C	C	T	ACCCTATAGCAGAACATTTG[T/C]ATTGAGTGGCACGTGGGCTG	CCAACCCTATAGCAGAACATTTG	CCCACGTGCCACTCAAT	1392	11
WIAF-3299	C/A	A	C	TGATAATCCCTCATGGAACC[C/A]AATTTCTGTATTTTAAAAAC	TGATAATCCCTCATGGAACC	AAGGTGTCGGTTTTTAAAATACAG	1394	10
WIAF-3300	G/C	C	G	TACAGAAAGAAAATTTTAGA[G/C]TAAACTTGGAACTTTGGATA	CAAATGGGTAAATGTACAGAAAGAA	TCTAGTTATATCCAAAGTTCCAAGTTT	1395	3
WIAF-3302	G/C	C	G	TTTTCTGCACTGTAGGCACA[G/C]TCTCTCAAGCATATCCTGAT	TTTTCTGCACTGTAGGCACA	TGATCAGGATATGCTTGAGAGA	1397	12
WIAF-3305	G/A	A	G	AGCCGGGGCGGTGCGATTCT[G/A]CCACACGCCACGCTCTACTA	GGGCGGTGCGATTCT	AGAGCGTGGCGTGTGG	1399	17
WIAF-3307	A/G	A	G	TATTTAAAAACTTGTTAAGA[A/G]GTACAAGCCAGTCTTCTGCT	GAGGTCCACATATTTAAAAACTTGTT	CAGCAGAAGACTGGCTTGTAC	19245	13
WIAF-3308	A/G	A	G	CTTCAAAGAAAATATAGCCT[A/G]AACAGCTATCAAATCCAAAC	TGCACTTCAAAGAAAATATAGCC	CAGAGTTTGGATTTGATAGCTGT	19246	13
WIAF-331	C/T	C	T	CATAGATTTTTTGACGACTA[C/T]GTGGCCATGCCATTCCTGTA	GGTTTGACCTGCATAGATTTTT	GGAATGGCATGGCCAC	2657	18
WIAF-3310	G/A	A	G	CAGGCCCGGGAGCCGGGGTC[G/A]TGGGAGCTGGGTTTCCCAGG	TGTTGGACATGCAGGCC	GGAAACCCAGCTCCCA	19172	9
WIAF-3312	G/A	A	G	CAAGGAAGAACAATAAAGTC[G/A]TAGGTTCAAATGTAAGAAGC	GGTAACCAAGGAAGAACAATAAAGTC	CATTGCTTCTTACATTTGAACCT	19601	6
WIAF-3321	T/A	A	T	CCCTAGTACTTGGGCCCTCA[T/A]CCTTCTCGTAGTACCTCTTG	CCTGGCTTTGTTTAGACCCT	CGCAAGAGGTACTACGAGAAGG	1401	7
WIAF-3323	G/A	A	G	ATCCCCATTCTGAAAGACCG[G/A]GAGCCTGGGAGGTGTGACCC	CCCCATTCTGAAAGACCG	TCACACCTCCCAGGCTC	1403	9

续表

WIAF命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi编号	NwLib chrom
WIAF-3324	A/G	A	G	TTACCTAGCAGAACATCCCC[A/G]TGGTTTGGGTGATTCTGAAT	GCTAATTACCTAGCAGAACATCCC	TTCAGAATCACCCAAACCA	1404	16
WIAF-3325	C/T	C	T	TCTTACTTTATAAGGAATAA[C/T]GTATGAATCATAAAAGAAGA	AGGAAAACAAAGTTCTTACTTTATAAGG	CGCTCATTCTTCTTTTATGATTC	1405	9
WIAF-3328	T/C	C	T	TTAAAAATATAATCTGGGTT[T/C]GGCTTAAGTATAGTGGGCCC	AACAATGTTTAAAAATATAATCTGGGTT	GGGCCCACTATACTTAAGCC	19602	6
WIAF-3330	G/A	A	G	TAGAAAGAAACACACTTTAA[G/A]AAATCAAAATTCTCAATTCA	AGCAATTTAGAAAGAAACACACTTT	GCCTGAATTGAGAATTTTGATTT	1409	12
WIAF-3336	T/C	C	T	GCATAGACTGTCTTATTTAT[T/C]GAAATACTCACAGAATGTAT	TGAGGGCATAGACTGTCTTATTT	TTTGCATACATTCTGTGAGTATTTC	19256	10
WIAF-3338	C/T	C	T	CATATGGTCCACTTGATAGA[C/T]GTTCAGGAGGGAAGAAATGC	AATTTCATATGGTCCACTTGATAGA	GCATTTCTTCCCTCCTGAAC	19271	2
WIAF-3339	A/G	A	G	TGAAGATACTTTATCTTAGG[A/G]TGACCTCTGTCTCTTTCCAC	TTACCAAATGAAGATACTTTATCTTAGG	CAGTGGAAAGAGACAGAGGTCA	19184	5
WIAF-3340	C/T	C	T	AAAATGCATTATGTCTAATA[C/T]TTCATGTCAAACAGCAGTAG	GGTGAAGGGACTTGCAAAAA	TGTCTACTGCTGTTTGACATGAA	19603	13
WIAF-3341	T/C	C	T	TGAAGCTATGAAGACTGGGG[T/C]GATGTCCACCTTGCTGGATC	TGAAGCTATGAAGACTGGGG	TCCAGCAAGGTGGACATC	19241	8
WIAF-3342	A/G	A	G	TGAGTGTCTTCCTAGAGTCC[A/G]TGGCTTTGGGTGTGTGAGAG	AAACTGAGTGTCTTCCTAGAGTCC	TCACACACCCAAAGCCA	1415	5
WIAF-3343	T/C	C	T	ATTGAATCTACAGCAAATCA[T/C]GATAATTTATTCCAAAAGTG	TGTGATTGAATCTACAGCAAATCA	CATCATCACTTTTGGAATAAATTATC	1416	17
WIAF-3344	C/T	C	T	CTCTCAGGCTAGAATGAGAG[C/T]GGGAATTTTTATTTATTTTG	CTGCTCTCAGGCTAGAATGAG	AGCGAGCAGTGAACAAAATAA	1417	6
WIAF-3345	C/G	C	G	GGGAATTGCTCTCCTCTATT[C/G]AGATCTTGTGATTCTGGTGC	GGGGAATTGCTCTCCTCTATT	TGCACCAGAATCACAAGATCT	1418	8
WIAF-3346	C/T	C	T	GCACCTTGCTCCACTTCTGG[C/T]TAATCATCACACTCTATCTG	CACCTTGCTCCACTTCTGG	TTCTGCAGATAGAGTGTGATGATT	1419	8
WIAF-3348	G/A	A	G	AGGGAGAGGGTGGCATGCTG[G/A]AGCCCATGCTGGTCACGATG	GAGAGGGTGGCATGCTG	CGTGACCAGCATGGGC	1420	2
WIAF-3354	T/C	C	T	AGACTCAGTTGTCCAACATT[T/C]CCCTGGGGTGAGCTGATGAC	GCAGACTCAGTTGTCCAACA	TCAGCTCACCCCAGGG	1426	6
WIAF-3355	T/G	G	T	GCAGAGAAAACGGAGTAACT[T/G]CAAGGCGCCATCCATCCTCT	TGGCAGAGAAAACGGAGTAA	ATGGATGGCGCCTTG	1427	6
WIAF-3358	T/C	C	T	TAATAACCCCTTAAACTCCA[T/C]GCTCTCTTCTATACACCACT	CCATAATAACCCCTTAAACTCCA	GAGGAGTGGTGTATAGAAGAGAGC	1430	9
WIAF-336	T/C	C	T	TTGAGTATTCTTGCTTTGAT[T/C]GTCTACGTAAGCATGTAAGA	CATTTATTGAGTATTCTTGCTTTGAT	GTTGTAGTCTTACATGCTTACGTAGAC	2660	15
WIAF-3361	T/C	C	T	CAGATGGCCATTAGAAGTTA[T/C]TAAACAAATATTTCTCCCAA	TTTACAGATGGCCATTAGAAGTT	CTTTTGAGTTTTGGGAGAAATATT	1433	4
WIAF-3365	G/T	G	T	CAAGCATGTGACAGTCTGAG[G/T]TCCCTCCAAGAGCCAGGGGA	TCAAGCATGTGACAGTCTGAG	CCTGGCTCTTGGAGGGA	19604	1
WIAF-3367	C/A	A	C	GATTTCACTGAGGAGCTGAC[C/A]TTGGTGAGGAGCTGTACTCC	GGATTTCACTGAGGAGCTGA	GGAGTACAGCTCCTCACCAA	1438	13
WIAF-337	A/C	A	C	TCTTCTAATTTTAAGTGAGA[A/C]TCTTTTAAACACCTGTTAAA	AATGGATCCATAAACTTTCTTCTAAT	GCTACATTAAATTTAACAGGTGTTTAAAA	2661	2
WIAF-3372	C/T	C	T	ATCTTCTAGGTCTTTCCCCA[C/T]GTGTGTCATGTAACAGGTGC	TAATCTTCTAGGTCTTTCCCCA	TGCACCTGTTACATGACACAC	19288	4
WIAF-3380	A/G	A	G	TGTCATTTCTGCTCACACTT[A/G]CCTCACTTATTAAATGGTAA	TGTGTCATTTCTGCTCACACTT	GCCCTTACCATTTAATAAGTGAGG	1449	3
WIAF-3383	C/G	C	G	CGCATGACAATGGAATGTCT[C/G]TTTAGTTACTTAGGTCTCTA	TGTAAAAATTGGAGTCCTTCAAC	GGAGAGAGAAAGTAGAGACCTAAGTAAC	19605	14
WIAF-3384	C/T	C	T	CCATTCTCTGCCCTAATCTC[C/T]CAAAGAGCATTCAGCTGAAT	TCCATTCTCTGCCCTAATCTC	GCAGCTTTTCATTCAGCTGA	1451	13
WIAF-3385	C/A	A	C	AATCTCTAGGGGAACCCTCA[C/A]ACACTACTGGGAGGAATGTA	AATCTCTAGGGGAACCCTCA	GGGAAACCTCCATACTGCTC	1452	3
WIAF-3386	T/G	G	T	TTTCTTAACAGCTTTAAAAT[T/G]CAGTATTTTATTTATTAACA	GGACAATTTCTTAACAGCTTTAAAA	TCTAGCGCATTTAATGTTAATAAATAAA	1453	13
WIAF-3387	C/T	C	T	CTTAACCTTTAGTTGTTGAA[C/T]GCCTACTTCGAGTAAGAGAT	CACTTACACTTAACCTTTAGTTGTTGA	CACAATCTCTTACTCGAAGTAGGC	1454	3
WIAF-3390	G/A	A	G	GACCAGCTGGTCAATTGACA[G/A]TCATCCTGGCAGGGGATCTG	ACCAGCTGGTCAATTGACA	CCCCTGCCAGGATGA	1457	7
WIAF-3392	G/A	A	G	ACTGTCCCATGCTGGTCAGA[G/A]TGGATAGAACTTAGGGAGTT	TGTCCCATGCTGGTCAGA	TCAGAACTCCCTAAGTTCTATCCA	19219	
WIAF-3393	G/T	G	T	CAGGTTTGAAATTCTTGGTT[G/T]GAGTTTCCTCTTCAAGAATG	GGTCAGGTTTGAAATTCTTGG	ATTTTCATCATTCTTGAAGAGGA	1459	19
WIAF-3394	G/A	A	G	CCCTTTGTACCTGATGTGCC[G/A]TTTCCTCTAGGCTGCCTTTA	CCCTTTGTACCTGATGTGCC	CCTTAAAGGCAGCCTAGAGGA	1460	19
WIAF-3396	T/C	C	T	CTCTAAGTTTTCCTGTCAGC[T/C]CTGACTCCTGTTTCCCACTC	TTTAAAAATGTTATAGCCACTCTAAGTTT	GAGTGGGAAACAGGAGTCAG	1462	10
WIAF-3402	C/G	C	G	GCATGGTGGTCGGCTTTCAG[C/G]AGCCGGGCACAGTTCACAGT	TGGTGGTCGGCTTTCAG	TGAACTGTGCCCGGCT	1467	5
WIAF-3404	A/G	A	G	GCTTCCAGCCCCAGCCCTTC[A/G]CAAACATTATTCTACTTTTT	CAGCCCCAGCCCTTC	AAAAGAGAAAGGTATAGGTACAAAAAGT	1469	
WIAF-3407	C/A	A	C	AGTTCTTCTTGGTGTCTAAT[C/A]TAACTCTTTCTTGCTATTGC	TGGAAGTTCTTCTTGGTGTCTAA	CCAGCAATAGCAAGAAAGAGTT	19294	3
WIAF-341	G/C	C	G	CCATGTTCTTTCACTCATCA[G/C]CCTTCTGATTTTGATTCCCT	GCCATGTTCTTTCACTCATCA	AAGGGAATCAAAATCAGAAGG	2664	8
WIAF-3410	G/A	A	G	CAGGGAAAATGGCTAAATCA[G/A]TTATGTCACACTCATTCAAT	CAGGGAAAATGGCTAAATCA	CCCATTGAATGAGTGTGACA	1473	17
WIAF-3412	C/T	C	T	CTGGGATGTTTAATGCAAAC[C/T]GTACATTCTCAGCAGAGCAC	CCTGGGATGTTTAATGCAAA	TTGTGCTCTGCTGAGAATGT	1475	4
WIAF-3415	A/G	A	G	AAAGGAGATCCAACATTAGG[A/G]CTAATAAAACTGAAATGACC	GGAAAGGAGATCCAACATTAGG	TGCAGAAGGTCATTTCAGTTTT	1477	1
WIAF-3416	A/G	A	G	CAAGGATGGGACCCTCCTCC[A/G]CAGCGTGGTGCCCTGCCTCG	GGATGGGACCCTCCTCC	CAGGGCACCACGCTG	19237	
WIAF-3418	T/C	C	T	GAGGTTGAATGGGGCACCAA[T/C]TGCTTGGCTGGAAGGAAGTA	TTGAATGGGGCACCAA	TCCTTCCAGCCAAGCA	1479	4
WIAF-3420	A/G	A	G	AAGGTCAGAGTGGGACCGGC[A/G]GCTAAATCATCAGGAGTGGA	TCAGAGTGGGACCGGC	ACTCCACTCCTGATGATTTAGC	1481	8
WIAF-3421	C/T	C	T	CTTGTGCTCAGCTTATGCCC[C/T]GCATTTATAAAGAGATTGTC	TGTGCTCAGCTTATGCCC	AGGAAAGACAATCTCTTTATAAATGC	1482	16
WIAF-3422	G/A	A	G	AGTGAAAAGACAACCAATCC[G/A]CTGAATAAGAGAGACTATTT	AAAGTGAAAAGACAACCAATCC	GTCCCTAATCAGATCTATAATTTGTAAATA	1483	19
WIAF-3425	T/C	C	T	CTTTCCTAACCATTCGAGGG[T/C]TTTGTCCATTTCCCAAAGGA	TTTCCTAACCATTCGAGGG	CCTTTGGGAAATGGACAAA	1486	16
WIAF-343	G/C	C	G	TCCCATGCTTGGATGTCACA[G/C]TTATGTCAAGTTAATATAAA	CCATGCTTGGATGTCACA	GCACTTAGAAATGTTTATATTAACTTGACA	2665	3
WIAF-3430	G/A	A	G	TAATAAGTGCATGTGTGCAT[G/A]TCTGTGGCAAGTGTGTGTGT	AAAATAATAAGTGCATGTGTGCAT	CACACACACTTGCCACAGA	1491	1
WIAF-3437	G/A	A	G	AAACAAAACAATATAAAGTC[G/A]TAAGAGATAACCCTGATAGG	GGTCAAAAGGAACCCCC	AGCCCTATCAGGGTTATCTCTT	1498	1
WIAF-3442	G/T	G	T	CAGAGGCCAAAGTCTGGGGG[G/T]TAGAAAGAAGGAAGCTTGAC	GGCCAAAGTCTGGGGG	GGGTCAAGCTTCCTTCTTTCT	1503	2
WIAF-3443	G/T	G	T	GACCCAAACAAAGTTTCATT[G/T]ACAGTATATACAGTCAGGCC	TTGACCCAAACAAAGTTTCAT	CCCAAGGCCTGACTGTATATAC	1504	2
WIAF-3444	G/A	A	G	CATACCTTTCGACAATGCCT[G/A]CAGCCTCATGTAAGGTTATA	CATACCTTTCGACAATGCCT	TGTGTATAACCTTACATGAGGCTG	1505	6
WIAF-3449	T/C	C	T	AGCTGCCTGGGTTCTTGAGT[T/C]TGAAATCCCTGTCTTAAAGG	GCTGCCTGGGTTCTTGAGT	CATCCTTTAAGACAGGGATTTCA	1510	2
WIAF-3452	T/C	C	T	ATTAAAAACAAGGTATGGGA[T/C]GCCTGATAACACTACAACTT	GGAGATTAAAAACAAGGTATGGGA	AAAAGAAGTTGTAGTGTTATCAGGC	1513	6

续表

WIAF 命名	等位基因	等位基因 A	B	芯片中的位置	上游引物	下游引物	ncbi 编号	NwLib chrom
WIAF-3455	T/G	G	T	CACTTTGAAGCAACCAGATG[T/G]GATGAGGACTCAATATCAGG	CACCACACACTTTGAAGCAA	TCTCCTGATATTGAGTCCTCATC	1516	3
WIAF-3459	C/T	C	T	ACCACTCTGGCCCTTTTGAA[C/T]GAACTTGGTCTGTGCAATCT	CACTCTGGCCCTTTTGAA	GAGATTGCACAGACCAAGTTC	1520	4
WIAF-3460	C/T	C	T	CTCATGACATTATAATGGAG[C/T]TGAAAAATTGCTATTGGCAG	CCATGATGGTCTCATGACATT	GAACTGCCAATAGCAATTTTTC	1521	1
WIAF-3462	A/G	A	G	CCTTGCTATGAAGCGGAAAT[A/G]AAATTAGGGTTCTGTTGGCA	CCTTGCTATGAAGCGGAAAT	TGCCAACAGAACCCTAATTT	1523	2
WIAF-3464	C/G	C	G	GCATCCATTCCTGTGCCTTT[C/G]GTGCTCCTGTATGTGGAAGT	CATCCATTCCTGTGCCTTT	TGACTTCCACATACAGGAGCA	1525	5
WIAF-3467	G/T	G	T	GGGGTCCTTACTTGTATTCT[G/T]TTATCAGCTGATTTTGAAAC	AATATGGGGTCCTTACTTGTATTCT	ATTATTATATGTTTCAAAATCAGCTGATA	1528	6
WIAF-3468	G/T	G	T	TGAGATGCAGAGAACGTACA[G/T]GATCCCATTTATTTAAAGCA	TGTGAGATGCAGAGAACGTACA	AGTTTGCTTTAAATAAATGGGATC	1529	2
WIAF-3469	G/C	C	G	AGAAACATAGGGAAATGATT[G/C]AGAGAGATAATAGAAGAGAA	GGTGATAGAAACATAGGGAAATGAT	TGTAATACTCATTTCCACTTTCTCTTC	1530	10
WIAF-3472	G/A	A	G	TCTCTAAAGATGGCTCTTAT[G/A]TTTGAGGAAACTGAGACTCA	GAACAATTCTCTAAAGATGGCTCTT	GCAAGTAAACCTCTTTGAGTCTCA	1533	9
WIAF-3473	C/A	A	C	AGACCACCAGATATCATGAG[C/A]CCCCTGGCAGAAATGCATGC	CAAAGAGACCACCAGATATCATG	GCATTTCTGCCAGGGG	1534	2
WIAF-3474	G/C	C	G	TCACTTTTGCCATCTTATCT[G/C]TTTAAAACTCTTTCAAGAAA	CAGTCACTTTTGCCATCTTATCT	CCTTTCTTTTCTTTTCTTGAAAGAG	1535	6
WIAF-3477	T/G	G	T	CGTACACTGGTACCACCAAA[T/G]GTTGCATCGTTGGGGGAAAG	TCTCACTGTGTGTTTAAGAGACG	CCCCCAACGATGCAAC	1538	9
WIAF-3479	C/A	A	C	ACGCTGAGCTGCAAGCAGAC[C/A]CCACAGCACCGAGATGGGGG	GCTGAGCTGCAAGCAGAC	CCATCTCGGTGCTGTGG	1540	21
WIAF-3483	G/C	C	G	GGGTGGTGGAGCTCCAGGGG[G/C]CTGCCCATGAAAGGGCTGGG	GGTGGAGCTCCAGGGG	GCCCTTTCATGGGCAG	1544	11
WIAF-3485	T/C	C	T	CTGGTGTGAATGATTCTCAG[T/C]TCAGGATTTCACGAAGCTGT	GGCCTGGTGTGAATGATTC	CAGCTTCGTGAAATCCTGA	1546	6
WIAF-3486	T/G	G	T	ACATGAAGGTATGCATAAAC[T/G]GTCTTCACATTTAATTTTGT	GCACATGAAGGTATGCATAAAC	CGAATCATACAAAATTAAATGTGAAGA	1547	21
WIAF-3492	T/C	C	T	TCTCTCTTTGTCCATGGCTT[T/C]CTGAGATTCAGTGTCCTCGA	TCTCTCTTTGTCCATGGCTT	CATCGAGGACACTGAATCTCAG	1553	11
WIAF-3494	A/C	A	C	TGCCAAAATTCTCTAGAAGA[A/C]AGAATTATTATGGCTTCCCA	TTTTGCCAAAATTCTCTAGAAGA	GAGTTACTGGGAAGCCATAATAAT	1555	16
WIAF-3496	T/C	C	T	ATGGATGGAACATTAACTGA[T/C]GTTTGTGTTGTGCGTAGGAA	GCATGGATGGAACATTAACTGA	TTCCTACGCACAACACAAAC	1557	21
WIAF-3497	A/G	A	G	AGGAGAGGTAGGTTAGGGAT[A/G]AAGACTGAGCTCATTAAAGT	AATAGTAGGAGAGGTAGGTTAGGGA	CCAAAACTTTAATGAGCTCAGTCTT	1558	21
WIAF-3498	G/A	A	G	AAAGTTTTGGTTGTCTTCAG[G/A]TCATCGAATCAAAAATATAT	CATTAAAGTTTTGGTTGTCTTCAG	TGACATTAATATATTTTTGATTCGATG	1559	21
WIAF-350	T/C	C	T	TCGTTAAGTGCTGGATATAC[T/C]TGGCTTGCACCGGACACCTT	CCTTCTCGTTAAGTGCTGGA	CCCTCCGTAAAAGGTGTCC	2670	7
WIAF-3500	A/G	A	G	GTTTTTTTCATCTAAGAAAT[A/G]AGAAACTCTGAGCAGCAAAT	CACATGAATATAACTCGATGAGTTTTT	CAACTATTTGCTGCTCAGAGTTT	1561	14
WIAF-3501	C/G	C	G	TGACGCAATTGTGTGACCTA[C/G]CTCTAGAGTCCTCTACCAGA	TGACGCAATTGTGTGACCTA	CTAACTTTCTGGTAGAGGACTCTAGAG	1562	12
WIAF-3504	G/C	C	G	GACACTGTTGACTTGGCTTG[G/C]GTAAAGGTACACATGAAAAC	AGACACTGTTGACTTGGCTTG	AAGCAGTTTTCATGTGTACCTTT	1565	16
WIAF-3505	T/C	C	T	GGTGCTATTGCTCAGGGTTA[T/C]GCTAACGTTTTAATCAGCAT	GGTGCTATTGCTCAGGGTTA	GCAATGCTGATTAAAACGTTAGC	1566	1
WIAF-3508	G/A	A	G	CAATAATTTGTTAGAATAAC[G/A]GTAAGTGCACAGGCTCAATA	TGATTTTCCACAATAATTTGTTAGAA	TTATTATTGAGCCTGTGCACTTAC	1569	21
WIAF-3511	G/A	A	G	TTGTGGCACAAATGGACGAC[G/A]TGTGCTTCTTAAGAAAGACC	TGGCACAAATGGACGAC	ACTGGTCTTTCTTAAGAAGCACA	1572	20
WIAF-3512	T/G	G	T	TTGAGGCCAGTTTCTCAGTG[T/G]TTAACTAGAACCTTAAAGAT	TGAGGCCAGTTTCTCAGTG	GGAGTTAACATCTTTAAGGTTCTAGTTAA	1573	12
WIAF-3518	G/A	A	G	ACAGCTGCCAAATACTGGGC[G/A]CGGGCCAGACTCTTCAGGCA	GCTGCCAAATACTGGGC	CTGAAGAGTCTGGCCCG	1579	9
WIAF-3520	T/C	C	T	CCCACATAGCTGTGTAGACA[T/C]TGATATGTCAGCAGAAACCT	GATACCCACATAGCTGTGTAGACA	CAAAGGTTTCTGCTGACATATCA	1581	9
WIAF-3527	T/C	C	T	ACGTCTTACAGAATACCTAT[T/C]GAGATATTACACCTATTTTA	GGCTACTATTACGTCTTACAGAATACC	CCTAGCTCTTTAAGATAAAATAGGTGTAA	1588	9
WIAF-3529	C/T	C	T	GTCCAATGGAATACAACATT[C/T]CAGAACAGGGTAAGCTGGGG	CCTGTCCAATGGAATACAACA	CCCAGCTTACCCTGTTCTG	1590	1
WIAF-3531	T/G	G	T	CTGGAGGTGTGGGCATTTGA[T/G]TGAAATTCAGACTGCCTTAA	GGAGGTGTGGGCATTTGA	ATTTAAGGCAGTCTGAATTTCA	1592	6
WIAF-3533	T/C	C	T	AGATAGTTATCAGCAGTCTT[T/C]GTGAAGTCTTCCAACTAATA	CAATCACAGATAGTTATCAGCAGTC	AATGAAATATTAGTTGGAAGACTTCAC	1594	2
WIAF-3534	A/T	A	T	TGGAAAGGAAGACAGCTATG[A/T]CCCAACACCCACCAAAAGAC	GGTGGAAAGGAAGACAGCTATG	TTTTGGTGGGTGTTGGG	1595	3
WIAF-3537	G/A	A	G	GATTCTGCGCTAAGTGCTCT[G/A]CTGTGAGACATCACAGGTTG	GATTCTGCGCTAAGTGCTCT	TGACAACCTGTGATGTCTCACA	1598	10
WIAF-3538	G/A	A	G	CTCCTTCTCCAGGGAATCCC[G/A]CTGATCAGGGACTGGGAAAA	CCTTCTCCAGGGAATCCC	ACCTTTTCCCAGTCCCTG	1599	2
WIAF-3540	G/C	C	G	AGCTGCTCATCTGGGGATTG[G/C]AGCTGGAGCATCTGTCAAGG	CTGCTCATCTGGGGATTG	CCTTGACAGATGCTCCAGCT	1601	6
WIAF-3542	G/A	A	G	AAGGCCTGAGAACCAAGAGA[G/A]TCAATGGTGTAAATTCTAGT	AGGCCTGAGAACCAAGAGA	GGACTTAGACTAGAATTTACACCATTGA	1603	10
WIAF-3543	A/G	A	G	GGTCTAGTACAAGATGGTAG[A/G]AGAGAGGGTTAGCCATCCTG	TCTTCTGGTCTAGTACAAGATGGTAG	CAGGATGGCTAACCCTCTCT	1604	13
WIAF-3546	A/C	A	C	ATCAGTTTGTACTGATATTT[A/C]AAGTGCACACATGGGTACAT	AAACATTGTTTATCAGTTTGTACTGATAT	TATATGTACCCATGTGTGCACTT	1607	19
WIAF-3547	T/A	A	T	GAATAATCTGACTGATGGAC[T/A]ATGCTCACAAACAGATTTTA	GCTGAATAATCTGACTGATGGAC	TGTCTAAAATCTGTTTGTGAGCA	1608	5
WIAF-3549	T/C	C	T	ATTATGGAAGGTCTTAAACA[T/C]CAGCCATGTTTTTAGGAGTT	CCAATTATGGAAGGTCTTAAACA	CAAACTCCTAAAAACATGGCTG	1610	12
WIAF-355	G/A	A	G	AAGAAAGAAGAGCCTGGGAA[G/A]AGGGAATGAGAAAAGCACAA	AAGAAAGAAGAGCCTGGGAA	TTGTGCTTTTCTCATTCCCT	2134	8
WIAF-3553	C/T	C	T	ACGTATGTTCACAAAAGCCC[C/T]AAGTGATTTTGAAGCAGCTG	ACGTATGTTCACAAAAGCCC	CCAGCTGCTTCAAAATCACTT	1614	5
WIAF-3554	A/C	A	C	GACAGGTCCATCATGCCCCA[A/C]TGGAAGACCCACACCCCTAA	GGTCCATCATGCCCCA	GGGGTGTGGGTCTTCCA	1615	9
WIAF-3556	C/A	A	C	AAGAAATGTCATATAAAACA[C/A]TAGAGAATGTAGTTCCTGGA	GCCAAAAATAAAAAGAAATGTCA	GATTGTCCAGGAACTACATTCTCT	1617	2
WIAF-3557	A/T	A	T	TGAAGAAAACTAGGAGTACT[A/T]TTATAGAAATGTCTTTTAGG	TTTTTCATGAAGAAAACTAGGAGTACT	TGAACCCCTAAAAGACATTTCTA	1618	4
WIAF-3558	T/C	C	T	AGGCACTGCAGGTACAGGGA[T/C]GGTGAGCTGTGAGACTGCTC	GCACTGCAGGTACAGGGA	GAGCAGTCTCACAGCTCACC	1619	2
WIAF-3559	A/T	A	T	AGAAATCATCTTCAAATCTC[A/T]TTCACTGGAGGAGGCCAAGA	ACAGTTTAGAAATCATCTTCAAATCTC	TGGCCTCCTCCAGTGAA	1620	5
WIAF-3562	C/T	C	T	AGGTGGTGTTTTAGATACTA[C/T]ACAGTAAGAATATAACAATG	ACACTGAGGTGGTGTTTTAGATACT	TACTTGTTCTATGCATTGTTATATTCTTAC	1623	21

续表

WIAF命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi编号	NwLib chrom
WIAF-3564	C/G	C	G	ATCAAAGTTGTAGACCTCTG[C/G]CATGCCCTGATGTAGAGTTT	GTAATAAAATCAAAGTTGTAGACCTCTG	CAAAACTCTACATCAGGGCATG	1625	12
WIAF-3565	A/G	A	G	ATGGAGGCTTCATTATGTAG[A/G]CATGATTGATTAAATTGTTG	CCTGAACTCAGCTCTTCTGG	AGTGGCCAACAATTTAATCAA	1626	1
WIAF-3566	T/C	C	T	CTCCAGTGAGACATGGATGA[T/C]ACCCACACTTCACTCTTGAT	CTCCAGTGAGACATGGATGA	GGATCAAGAGTGAAGTGTGGG	1627	3
WIAF-3568	A/C	A	C	CACCTCCATATCCTGTCCCA[A/C]TGGAAGGTCTTCAGGGGTAA	ACCTCCATATCCTGTCCCA	AGGCATTGTATTCATAGATGACAG	1629	3
WIAF-357	C/T	C	T	GAGGGTCTGTGTGAAGGGGC[C/T]GGCTTCTCTTGGTGCCTGCT	GGTCTGTGTGAAGGGGC	CAGGCACCAAGAGAAGCC	2675	8
WIAF-3572	G/A	A	G	CCGGCACATTACAGGTCTTC[G/A]CGGGAGGGCAGCTGTGTCGG	CCGGCACATTACAGGTCTTC	ACAGCTGCCCTCCCG	1633	22
WIAF-3573	T/C	C	T	TTATTGTGGTCTTAATTCTA[T/C]ACATAATTACATTAATTACC	CTCCCTTATTGTGGTCTTAATTCTA	TTCTCTCATGGTAATTAATGTAATTATGT	1634	8
WIAF-3574	C/T	C	T	TCAGATTTATTGATTTCCAT[C/T]GTATTGACTTTAAAAATGGT	GGAATTCAGATTTATTGATTTCCA	GCTTCATACCATTTTTAAAGTCAAT	1635	16
WIAF-3575	G/A	A	G	GTTTCAATAGGACCAAGAAC[G/A]TTAGAGAATAAAGATCTTAG	CAATGTTTCAATAGGACCAAGAA	CATTTTCCATCTAAGATCTTTATTCTC	1636	5
WIAF-3580	T/G	G	T	TGGGTTAGGAAAAGTACAGG[T/G]ACAACCTTGACTTGCCTGAC	TTTGGGTTAGGAAAAGTACAGG	GTCAGGCAAGTCAAGGTTGT	1641	1
WIAF-3581	C/T	C	T	ACATGAAGTTGATATTGGTG[C/T]TTATAAATCACTCTCTCCCA	CCACATGAAGTTGATATTGGTG	GGACTGGGAGAGAGTGATTTATAA	1642	12
WIAF-3583	T/C	C	T	CACAAAGACCTCCCATCCTC[T/C]TGTTCCATCCACCTCTTCCA	CACAAAGACCTCCCATCCTC	GGAAGAGGTGGATGGAACA	1644	1
WIAF-3585	A/G	A	G	TTCTGTGGAATTTTGTGCTT[A/G]CTACTGTATAGTGCATGTGG	CATTCTGTGGAATTTTGTGCT	CCTACACCACATGCACTATACAGT	1646	7
WIAF-359	C/T	C	T	TGGTGCTGAACAAAACTGAA[C/T]GCTGTGCTTATCTTTCCTGA	TGGTGCTGAACAAAACTGAA	AATCAGGAAAGATAAGCACAGC	2677	5
WIAF-3594	A/G	A	G	AGCCACACGCATTACCCGAC[A/G]CAGCCCTCCTGACATCGGCA	CCACACGCATTACCCGAC	CGATGTCAGGAGGGCTG	1655	13
WIAF-3600	A/C	A	C	TCCACAGTCCTTCCATTTCA[A/C]CTTAGAGATGAACCTTTAAT	TCCACAGTCCTTCCATTTCA	ACCTTCATTGTATATTTTATTAAAGGTTC	1661	9
WIAF-3602	A/G	A	G	CAGGAAGAACCAAATATAGC[A/G]CAAGGTGTGGAAGGAGGGAC	CCTCAGGAAGAACCAAATATAGC	CCCTCCTTCCACACCTTG	1663	2
WIAF-3603	T/G	G	T	CTCATCCCTCCTACTTCAGA[T/G]TTATTACTATGGGGGCTGGA	CCTCATCCCTCCTACTTCAGA	TCCAGCCCCCATAGTAATAA	1664	7
WIAF-3607	C/T	C	T	CTCTGGGGTACCTGGGATTA[C/T]AGGCACGTGCCACGATGCCC	TCTGGGGTACCTGGGATTA	TCGTGGCACGTGCCT	1668	19
WIAF-3612	T/C	C	T	AGTTCTCTCTTACATAAAAG[T/C]CTGAAGTTATGCAGGCTAGG	GAGGTATGCTTAGTTCTCTCTTACATAA	TCCTAGCCTGCATAACTTCAG	1673	3
WIAF-3613	G/A	A	G	CAGGAAGCTATGCTAGGCAC[G/A]GCACGGGTTGACTGAAAATG	CAGGAAGCTATGCTAGGCAC	TTTTCAGTCAACCCGTGC	1674	6
WIAF-3614	C/T	C	T	TTTTTCAGGCTTTTTGAGTT[C/T]TGTGGACAGCTCCATTACTC	CCAGGCTCTTATTTTACTTAGTTTTT	GGAGTAATGGAGCTGTCCACA	1675	10
WIAF-3616	C/T	C	T	CTGAATTGACTTTTCCTCTC[C/T]GCTTCTGTGGGCGAACTCCC	TGCCTGAATTGACTTTTCCTC	AGTTCGCCCACAGAAGC	1677	2
WIAF-3618	G/C	C	G	TGTGCAAAATGCTAAGCAGG[G/C]GTTTCACTCAGGGCAGCCCG	GTGCAAAATGCTAAGCAGG	GGCTGCCCTGAGTGAAAC	1679	22
WIAF-3619	G/A	A	G	TGATGAGTTTGATTTGAAAC[G/A]CCACTATTGAGCACTCAATC	GGAGATGATGAGTTTGATTTGAAA	CAGATTGAGTGCTCAATAGTGG	1680	15
WIAF-362	G/A	A	G	ACTGCACCCACTGACGAGAC[G/A]CAGAGACCTTGGACTACAGA	TGCACCCACTGACGAGAC	TCATCTGTAGTCCAAGGTCTCTG	2679	15
WIAF-3621	T/G	G	T	TGTTTAATTTTTAAATTATT[T/G]ATTTGTTCTGTAGGACAAGG	CATGCTGGCTAATTTGTTTTT	GAGACCTTGTCCTACAGAACAAA	1682	8
WIAF-3624	G/A	A	G	AAGACCAGGGTCATCCGTGG[G/A]AGCAGATGAGTAGGACACGC	CCAGGGTCATCCGTGG	GCGTGTCCTACTCATCTGCT	1685	12
WIAF-3627	A/G	A	G	TTACCTCATTAGAGGTTGCA[A/G]AAATGATAGATGCAGCCTAT	CTTTTACCTCATTAGAGGTTGCA	GGATAGGCTGCATCTATCATTT	1688	7
WIAF-3630	A/C	A	C	TTTTTTCCTTTCTTCTAATA[A/C]GGAGACTACACTCTAGATAA	TCTTTGGCCTCTGCTTTTT	CCTGTCTGTTTTGTTTATCTAGAGTG	1691	10
WIAF-3631	G/A	A	G	AATTTGCAAGGCTTTGCTAT[G/A]GTTACCCACATGCTAAAAAC	TTAATTTGCAAGGCTTTGCT	GGAGTTTTTAGCATGTGGGTAA	1692	12
WIAF-3632	T/C	C	T	GAGGAGCACGTGAAGGTAGT[T/C]ACAGCCGCTCAATTATAATA	AGAGGAGCACGTGAAGGTAG	CTGTATTATAATTGAGCGGCTGT	1693	1
WIAF-3637	A/G	A	G	GGTTTGAATTGGTGATTAGA[A/G]TAAAGATGCATATTCCCCTT	AGGGTTTGAATTGGTGATTAGA	GAAAGGGGAATATGCATCTTT	1698	10
WIAF-364	A/C	A	C	GGAAGTTAAATACTGATAGA[A/C]GATGCAAATTTGTCCTTTCA	AGCAAAAGGAAGTTAAATACTGATAGA	CATGAAAGGACAAATTTGCATC	2681	12
WIAF-3644	T/C	C	T	TATATACTTTGTCTTGTGAC[T/C]AACTAGAGCTTACAGGAGTA	GAAAATGTTACTTCATTACAAAGTGTTAC	CAGTTAACTTGTTTATACTCCTGTAAGC	1705	12
WIAF-3645	C/A	A	C	TAAACAAGTTTGTTGAACAG[C/A]TTTCTTTTCTAATATAATGA	CAAATTAAACAAGTTTGTTGAACAG	CAGTTGTCCATGTACACATAGTCA	1706	6
WIAF-3649	T/C	C	T	CACATTAAAAAATGAAGACA[T/C]GATCAAGGAGATGTAAGAGA	GACAGGTACAGAGCACATTAAAAA	ATTTGTCTCTTACATCTCCTTGATC	1710	11
WIAF-3650	T/C	C	T	AAGCAATGGCATGAGTTGCA[T/C]CAACATAATCAGCTCTTGAG	GCAATGGCATGAGTTGCA	TCCTCAAGAGCTGATTATGTTG	1711	7
WIAF-3651	C/T	C	T	TCCTCTATAAATGTTAATGT[C/T]TTAGCCAATGACATTTTTCT	TGCACTTTTGCATTCCTCTA	TTATCTAGAAAAATGTCATTGGCTAA	1712	15
WIAF-3653	C/T	C	T	GTACACAAACACAGGCCCAA[C/T]TGGATGGATACCCTAGCGCT	CACAAACACAGGCCCAA	GCGCTAGGGTATCCATCCA	1714	11
WIAF-3656	C/A	A	C	TTCTGAGATCTCATCACATA[C/A]CTGATCCCATTTAACTTCAG	AACACTTCTGAGATCTCATCACATA	ATAGACTGAAGTTAAATGGGATCAG	1717	7
WIAF-3659	A/G	A	G	AGGACCTCCGCTGCAAATAC[A/G]TCTCCCTCATCTACACCAAC	GGACCTCCGCTGCAAATAC	CATAGTTGGTGTAGATGAGGGAGA	1720	11
WIAF-3660	C/T	C	T	GTGGTGGTTATCTTTAAATA[C/T]GTCATGCTTTTTTTCTTAAC	TTACGGTGGTGGTTATCTTTAAA	CCACACTAAAAATGGTTAAGAAAAA	1721	17
WIAF-3661	A/G	A	G	TACACCCATGTGCTTCATCC[A/G]TATCCAACCCAGATCACCCA	CACCCATGTGCTTCATCC	GGGTGATCTGGGTTGGATA	1722	3
WIAF-3663	C/T	C	T	TGACTGGCCCAGTACCCTGA[C/T]TTCCAGAACCTGTAGCCGTC	CTGGCCCAGTACCCTGA	CGGCTACAGGTTCTGGAA	1724	2
WIAF-3664	G/A	A	G	TCTGGACCTCAGTTTCCTCA[G/A]TGAGCTGGTAAGAATGCACT	TGGCCCTGTCTCTTCTTTTT	TTAGTGCATTCTTACCAGCTCA	1725	3
WIAF-3667	C/T	C	T	GCCTGGGCAGGGACAACTCA[C/T]TTGTTATGATCAGAGGGAAA	TGGGCAGGGACAACTCA	CCTTTCCCTCTGATCATAACAA	1728	2
WIAF-3673	A/G	A	G	CTGAACTGTCATGAATTGGA[A/G]TTTTCTAGTAAAACACAATG	AGCTCTGAACTGTCATGAATTG	GCCAACATTGTGTTTTACTAGAAA	1734	3
WIAF-3675	C/T	C	T	AGAGATGCCCTGACACCCTC[C/T]AAGGTTCTACAAGGTGACCA	GATGCCCTGACACCCTC	TTTGGTCACCTTGTAGAACCTT	1736	
WIAF-3676	C/T	C	T	AAGCTTTAGAAAGGTGACAA[C/T]ACATTTAATAGACTCCAGAA	CTTGGGAAATACACACAATGAA	GAAATATCAACTTCTGGAGTCTATTAAA	1737	6
WIAF-3677	C/T	C	T	TATCAATTGAGTAAGAGGAA[C/T]ATAAAATGTCTGAAATGCTA	TGTTGGTATCAATTGAGTAAGAGGA	CATCTGTTAGCATTTCAGACATTT	1738	6
WIAF-3679	T/C	C	T	TTAATCTGCTTTAGTTGAAA[T/C]AGCCAATACCAAAAGGGTAC	ATTTCAGGTTAATCTGCTTTAGTTG	TGTACCCTTTTGGTATTGGCT	1740	19
WIAF-3681	C/G	C	G	AATTACTTCATATTTGACAT[C/G]TATTTTGAAATTCTGTTTTA	GGAAGTCTGAATTACTTCATATTTGA	CCCTGTAAAACAGAATTTCAAAA	1742	8
WIAF-3683	C/A	A	C	CACTACAGGTAGCCTCTGCA[C/A]ATCGTTTAACCTCTCGGAGC	TCACTACAGGTAGCCTCTGCA	GCTCCGAGAGGTTAAACGAT	1744	15
WIAF-3684	G/A	A	G	GGAAATAAAATTGGAACGAA[G/A]AAAGGTTAGGAGAGTAGGGA	CAGGAAATAAAATTGGAACGAA	CCTTCCCTACTCTCCTAACCTTT	1745	20
WIAF-3688	T/C	C	T	ACACTGATCATGTGTCATAA[T/C]CACGTTCTATGGATTAAGCT	TCCACACTGATCATGTGTCAT	TGAGCTTAATCCATAGAACGTG	1749	9

续表

WIAF 命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi 编号	NwLib chrom
WIAF-3691	G/T	G	T	TTTTGGAAGTAATACTCTGC[G/T]TCTTCATACTGATCTGCTAG	TTCCTTTTGGAAGTAATACTCTGC	TGCTCTAGCAGATCAGTATGAAGA	1752	10
WIAF-3699	C/T	C	T	CAGAAAACCAAGGCTCAACC[C/T]AGTAGCGTCCTAGAGTAACG	CAGAAAACCAAGGCTCAACC	CCCGTTACTCTAGGACGCTA	1760	15
WIAF-370	A/G	A	G	GACACAATGCTGCCAATGCA[A/G]TTAGTATATAGAAATAATAC	CAATGCTGCCAATGCA	TAACAGCTGCGTATTATTTCTATATACTAA	2687	8
WIAF-3700	A/G	A	G	AACACACCTGAAGAAGACTT[A/G]GTTCCTGGTAACATTTCCTT	TTGAACACACCTGAAGAAGACTT	GCAAGGAAATGTTACCAGGAA	1761	6
WIAF-3704	C/T	C	T	ACCCCAAGTGAGTTCATTTA[C/T]GGCCCTGAGCCTTAGTTTCC	TGACCCCAAGTGAGTTCATT	GAAACTAAGGCTCAGGGCC	1765	6
WIAF-3705	C/G	C	G	ATGGTAACTTAGGGAAAAGT[C/G]GTAGTAGGACAGCAAACATG	TTTTGTGAATAAAATGTGTTTATGG	CCACATGTTTGCTGTCCTACT	1766	16
WIAF-3706	T/A	A	T	AGTCCCAGCGATGGGTTGGA[T/A]AAGTCTTGGGATTTCATGGG	CCAGCGATGGGTTGGA	CCCATGAAATCCCAAGACTT	1767	19
WIAF-3709	G/A	A	G	AGACGTTCACTGGGGGTCTT[G/A]GAACGTATGTCCCATGGGTA	GACGTTCACTGGGGGTCT	CCCATGGGACATACGTTC	1770	3
WIAF-3711	T/C	C	T	TTCTTTCCTGGATTCTTTGC[T/C]TTGCATAATTATACCTACAA	TTCTTTCCTGGATTCTTTGC	GCTGGTTTCATTGTAGGTATAATTATG	1772	10
WIAF-3712	A/G	A	G	ATCCACTTGGAGGGGTAGAT[A/G]TCGGTGGGGTTGGAGCGGCT	ATCCACTTGGAGGGGTAGAT	CTCCAACCCCACCGA	1773	
WIAF-3713	C/T	C	T	GATCAGGAGACTCTGTTAAT[C/T]ATTGATTCCCTTGGCATTCA	AAGATTGATCAGGAGACTCTGTTAA	GAATGCCAAGGGAATCAAT	1774	14
WIAF-3714	G/A	A	G	AGGGCTTGGGTCACAGCCTC[G/A]GGACTCATGGCTGCTCTGTA	GCTTGGGTCACAGCCTC	CAGAGCAGCCATGAGTCC	1775	16
WIAF-3720	A/G	A	G	CTTGCGCCTGCTCTCGCTCC[A/G]TTCTCAGTGTGGGCAGGAAC	CGCCTGCTCTCGCTCC	TCCTGCCCACACTGAGAA	1781	6
WIAF-3723	A/T	A	T	CCATAAGTGTAAGACTGGAT[A/T]GATGGAATAGGGATTACATA	CCAACCATAAGTGTAAGACTGGA	CTGATTATGTAATCCCTATTCCATC	1784	6
WIAF-3724	A/G	A	G	CAACACAGAGCAGCATGAAA[A/G]TCATTTTGTGAGAAACTGCA	CTACAACACAGAGCAGCATGA	TTTGCAGTTTCTCACAAAATGA	1785	8
WIAF-3725	G/A	A	G	CTAGGGTCCCGATGAGCTGC[G/A]TGTAGGTCGCTGGGGGGTGA	GGTCCCGATGAGCTGC	CCCCCAGCGACCTACA	1786	20
WIAF-373	G/A	A	G	GCCAAAACTAGGCCTCAGGT[G/A]CCCATTAAGCATGCTGTGAA	GCCAAAACTAGGCCTCAGGT	TCACAGCATGCTTAATGGG	2688	4
WIAF-3733	A/G	A	G	AAAGGGGTGGTATCATCTGA[A/G]GACTATGTACACAATTTGTG	GAAAGGGGTGGTATCATCTGA	CCACCACAAATTGTGTACATAGTC	1794	3
WIAF-3736	G/A	A	G	GACGGGACTGGCTTTTGGAC[G/A]TTGACATATTTTGAAAAGTG	GGGACGGGACTGGCTTT	AGGCACTTTTCAAAATATGTCAA	1797	9
WIAF-3737	G/C	C	G	ATCCCCTAACTTTATTATTC[G/C]TCATGGTCCTTACCAGCGGA	CACTCACTATCCCCTAACTTTATTATTC	CGCTGGTAAGGACCATGA	1798	X
WIAF-3738	T/C	C	T	CATAATAATATATAACATAC[T/C]ACAGGTCACATGTACATTTT	GGCATTTAACATTTAGCCATAATAA	AATGAAAATGTACATGTGACCTGT	1799	12
WIAF-3745	T/C	C	T	TCCTCACTGTTAGCGATCTC[T/C]GGAATGGCACAGAGTTGCTT	CCTCCTCACTGTTAGCGATCTC	GCAACTCTGTGCCATTCC	1806	11
WIAF-3746	T/C	C	T	ATTGTCTTCAGAACAATCAG[T/C]ACTACACTGCTTCCCATTAA	CAATATTATTGTCTTCAGAACAATCAG	GGTTTTAATGGGAAGCAGTGT	1807	6
WIAF-375	A/G	A	G	TGGGACTCTTGATCCAGAGA[A/G]GACAAAGCTCCTCAGTGAGC	TGGGACTCTTGATCCAGAGA	GCTCACTGAGGAGCTTTGTC	2690	12
WIAF-3750	T/C	C	T	GACAAGGAACAGCAGAATAG[T/C]GAACAAGTTTGTGGCTAAAC	GCTGACAAGGAACAGCAGAA	AATGTTTAGCCACAAACTTGTTC	1811	4
WIAF-3755	C/A	A	C	CTCATGGGTTGTAGAGCTGG[C/A]AGGGGACAGTGTTGTGTGTG	TCATGGGTTGTAGAGCTGG	CACACACAACACTGTCCCCT	1816	
WIAF-3756	T/C	C	T	TGTAAGTCCCTCCGATGCCT[T/C]CTTTCTGCTCCCCGCGCAGT	AAGTCCCTCCGATGCCT	CGCGGGGAGCAGAAA	1817	13
WIAF-3761	T/C	C	T	CCTCAGTAATTAGAAAATGA[T/C]ACAAATGCCGTCTTCACACT	AACTGCCTCAGTAATTAGAAAATGA	GTGTGAAGACGGCATTTGT	1822	10
WIAF-3762	G/C	C	G	ATACCACAGTAGACTTTTTG[G/C]CTATTTTCAAGCGGAATAAG	GTTTATTACAAATACCACAGTAGACTTTTT	TGTCTTATTCCGCTTGAAAATAG	1823	1
WIAF-3765	G/C	C	G	ACCTTTCACTTCCCTTCAAA[G/C]TTGTTTAAATTAGAGAATTC	AACCTTTCACTTCCCTTCAAA	CAACTAACTGAATTCTCTAATTTAAACAA	1826	4
WIAF-3766	T/C	C	T	TAATTTCTGCCCGCACCGCT[T/C]TGTGCCGAATGCACCAGTGC	TTTCTGCCCGCACCG	TGGTGCATTCGGCACA	1827	6
WIAF-3768	T/C	C	T	GAAAAACTCCCAGACATTCA[T/C]GACTTCATCCACCCTGCCTG	CAGAAAAACTCCCAGACATTCA	GGCAGGGTGGATGAAGTC	1829	16
WIAF-3769	A/C	A	C	GTGAAGGGTATTTTACGGGA[A/C]CTCTATAAAGCAGGAAGAAG	GGTGAAGGGTATTTTACGGGA	AAACTTGCTTCTTCCTGCTTT	1830	11
WIAF-377	C/T	C	T	AAGTGATCTACTCTCTTTAC[C/T]AAGTGTTACTTTGCATATAT	CCACACAGAAGTGATCTACTCTCTTT	CCCATAAAATATATGCAAAGTAACAC	2137	12
WIAF-3770	G/A	A	G	CTCTCAACCATGATGTGGTC[G/A]TATATTTATGGGTCAACATG	GCTCTCAACCATGATGTGGT	CAACACATGTTGACCCATAAAT	1831	11
WIAF-3775	G/A	A	G	ACAGAAACCACGTGAATGTC[G/A]GGAATATGGAAAGACCCTTT	GACAGAAACCACGTGAATGTC	CAAAGGGTCTTTCCATATTCC	1836	19
WIAF-3777	C/T	C	T	AAAATGCCTGGAAATGGATG[C/T]GAGAAAGACTGAGTCCCCAT	AAATGCCTGGAAATGGATG	CATGGGGACTCAGTCTTTCTC	1838	14
WIAF-3779	T/C	C	T	GATTGCCAAAGCGTCAAAGC[T/C]GGAATCAGGGTCTCTCAGTC	TGCCAAAGCGTCAAAGC	AGAATTGAAGACTGAGAGACCCT	1840	17
WIAF-3780	A/G	A	G	CTTCAATTCTCACTTTTCCC[A/G]CTTCTAATTCACCAGTGCTA	AGTCTTCAATTCTCACTTTTCCC	TGAGTAGCACTGGTGAATTAGAAG	1841	17
WIAF-3784	T/A	A	T	GGAACCCTCCCCAAAACCTT[T/A]CCCCAGACACATTCTCCTGT	AACCCTCCCCAAAACCTT	CAGGAGAATGTGTCTGGGG	1845	6
WIAF-3785	C/T	C	T	AGGAAGGAGCAGAAGTTGGA[C/T]CTGAAGGCAGGAAGCCAGCT	GGAAGGAGCAGAAGTTGGA	TGGCTTCCTGCCTTCAG	1846	2
WIAF-3788	A/G	A	G	ATTTATACCCCAAAACAATT[A/G]TGACAGTTAATAGGAAATCC	TGTCTCTATTTATACCCCAAAACAA	AAAGTAAGGATTTCCTATTAACTGTCA	1849	5
WIAF-3797	C/A	A	C	TAAGAATGCAAACACACAGA[C/A]AAATGCAGCTATCACAGGGG	CAAATAAGAATGCAAACACACAGA	CCCCTGTGATAGCTGCATTT	1858	10
WIAF-3798	A/C	A	C	TCTGTCACTGCAGGGCAAGA[A/C]ACCAGTTACACCTGAGCCTA	TGTCACTGCAGGGCAAGA	TCTAGGCTCAGGTGTAACTGG	1859	1
WIAF-3799	T/A	A	T	CATAGACAACTGGGATTTCC[T/A]GATATTAGGGAAATAAATTA	CACATAGACAACTGGGATTTCC	TCAATACTAATGCTCTAATAATGTAAATTG	1860	6
WIAF-380	C/T	C	T	GGGAAACCTCAGGTAGCTCC[C/T]GAAGATCTGTGCTTTCCAAC	GGGAAACCTCAGGTAGCTCC	TGTTGGAAAGCACAGATCTTC	2694	11
WIAF-3800	G/A	A	G	GTCTATTGATAGATAAACCC[G/A]CTGTGTCCTCACGGCTATCG	TTGGTGTCTATTGATAGATAAACCC	CGATAGCCGTGAGGACACAG	1861	1
WIAF-3801	C/T	C	T	TGGGATGAGCAATCCTGTTA[C/T]AGTACATCTGGACCTTCCCT	TGGGATGAGCAATCCTGTTA	GGTAGGGAAGGTCCAGATGT	1862	9
WIAF-3805	T/C	C	T	TTCTGGTGTCCATGGCAACA[T/C]GGCCTTATGGCTGTAATAGT	TGGTGTCCATGGCAACA	CCACTATTACAGCCATAAGGCC	1866	6
WIAF-3806	T/C	C	T	GATTTATTATTAAATCATTA[T/C]AGCAGTGAAGTGGTGCCATT	GATGATAAGTCTGCTTGATTTATTATTAAA	TGGCACCACTTCACTGCT	1867	6
WIAF-381	A/T	A	T	TTTTCCAGCCCCTGAGGAAA[A/T]GGAATGAACCACTCCCTGCC	CCAGCCCCTGAGGAAA	GCAGGGAGTGGTTCATTCC	2695	11
WIAF-3813	G/T	G	T	ACCAACCAGATTGGTTTAGA[G/T]GCCCATTAATCAAATCAAGT	CACCAACCAGATTGGTTTAGA	TGACTTGATTTGATTAATGGGC	1874	9

续表

WIAF 命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi 编号	NwLib chrom
WIAF-3818	A/G	A	G	GCTGCTCAGGCAAACCATGG[A/G]GCAAGATTGGAAATAATAAT	GCTCAGGCAAACCATGG	GACCATTATTATTTCCAATCTTGC	1879	4
WIAF-3819	G/A	A	G	CACCTTTTCCAGCTTCATAG[G/A]CCCCTTCCTCCAACTTCAAA	GGGGAAAATCTGTTTCCTCA	CTAATCTTGCTGCCTTTGAAGT	1880	20
WIAF-3821	G/T	G	T	ATTGGAAATAATAATGGTCT[G/T]CATACATCCTCCCATAGGGC	GCAAGATTGGAAATAATAATGGTC	GCCCTATGGGAGGATGTATG	1882	4
WIAF-3824	G/A	A	G	TCTTTCACAAAACAGCATTC[G/A]AAGGAGAAGGGAAAGTTCGC	TTTCTTTCACAAAACAGCATTC	GCGAACTTTCCCTTCTCCT	1885	22
WIAF-3825	G/T	G	T	AGAGGAGGGCACGGAATCTG[G/T]ATGTGCTCAGGAGGATGATA	GGAGGGCACGGAATCTG	AGTATCATCCTCCTGAGCACAT	1886	1
WIAF-3827	A/T	A	T	TGGGGGCTCCAGGAGTCACA[A/T]GTTTACATTGCAGAGAAGGA	CCAAGGACCAGAAGAAACTG	TCATTTCCTTCTCTGCAATGT	1888	3
WIAF-3828	T/C	C	T	CTCTATGGGTGACCTTGGCA[T/C]AGGTTCTTCTAAGAACTGGC	CTATGGGTGACCTTGGCA	AGAGCCAGTTCTTAGAAGAACCT	1889	16
WIAF-383	A/C	A	C	CAGATACAAGGAAATAAAAA[A/C]CACTTTTAGGAGATGAAAAC	AAAAGACTACAGATACAAGGAAATAAAAA	TTGTGTTTTCATCTCCTAAAAGTG	2697	2
WIAF-3831	G/A	A	G	CACCTGTTTCCAAACTTCAA[G/A]GTTAGGGCTTTGAAACATTC	CCACCTGTTTCCAAACTTCAA	GGAATGTTTCAAAGCCCTAAC	1892	9
WIAF-3834	G/T	G	T	TCTTGCTTAATGTTTTTACT[G/T]TTAATAGAAATAGAACTGAT	GGTAGTCTTATATCTTGCTTAATGTTTTT	GCCATAATCTTTATACCTATCAGTTCTA	1895	8
WIAF-3838	G/A	A	G	ATTGCCGAAGAAGGGGGAAA[G/A]AACTCCGAGTCCACTGTGGA	GCCGAAGAAGGGGGAAA	CCACAGTGGACTCGGAGTT	1899	11
WIAF-3839	C/T	C	T	GCTTCTTCATACATGCTGAG[C/T]GCCCTAGAAATGGAGGCTGT	ATGCTTCTTCATACATGCTGAG	CAGCCTCCATTTCTAGGGC	1900	11
WIAF-3840	A/G	A	G	GCCTCGGCCTCCTATCTGGA[A/G]AGTGGGACTGAGGGAGGATT	TCGGCCTCCTATCTGGA	GGGTGAATCCTCCCTCAGTC	1901	11
WIAF-3843	C/T	C	T	CATTATCTGAGACAACATAT[C/T]AAGTGAGCAGCAGTACAAAG	AACCTTTCATTATCTGAGACAACA	CCTTTGTACTGCTGCTCACTT	1904	6
WIAF-3855	G/C	C	G	ACCCTGGCAAGGCAAGGGGC[G/C]GCCATAGCACTCTTACCTGA	TGGCAAGGCAAGGGG	GGTCAGGTAAGAGTGCTATGGC	1916	16
WIAF-3858	A/G	A	G	ATAGCTACAAAGTTACAAAT[A/G]TATATTATGAGATTATTACA	TCAAGTGATTTCAAAGATAGCTACA	TGAAACTGAATATCATTGTAATAATCTCA	1919	4
WIAF-3861	G/A	A	G	GCCTGGGAGGATCTGGTGAT[G/A]GCAGAGAAAAGCAGGTGGGT	CCTGGGAGGATCTGGTGAT	CCACCTGCTTTTCTCTGC	1922	2
WIAF-3863	A/G	A	G	TTTGGTTCACAAATTGCTGC[A/G]TGGTACTTATATTCTCAGAA	TTGGTTCACAAATTGCTGC	GGCTTAGAGAATTTTCTGAGAATATAA	1924	15
WIAF-3864	T/C	C	T	TATGCTAAACAGGGGTGGAT[T/C]ATTCATGAGTTTTCCAGGAA	CCTATGCTAAACAGGGGTGG	CCTTTCCTGGAAAACTCATG	1925	22
WIAF-3865	C/T	C	T	CCCACAAAGTAATCTGACGG[C/T]GGGGCAAACTCAGGATCTCT	CCCACAAAGTAATCTGACGG	AGATCCTGAGTTTGCCCC	1926	3
WIAF-3873	G/A	A	G	GTGCTTTTGCTTGTTCCTTA[G/A]GAGGGGATTTAGTAGATGCT	GGTGCTTTTGCTTGTTCCTT	AAAGTAGCATCTACTAAATCCCCTC	1934	19
WIAF-3876	A/G	A	G	GGAAGGATATAAATGAAGAT[A/G]GGCAGACTGACAGCTGGTGG	TCCAGGAAGGATATAAATGAAGAT	ACCAGCTGTCAGTCTGCC	1937	15
WIAF-3877	G/A	A	G	CATGTAAGATGTCATGGAAG[G/A]AGTGGAATTGGATCTGAGTC	AAATCATGTAAGATGTCATGGAAG	CAGACTCAGATCCAATTCCACT	1938	6
WIAF-3880	G/C	C	G	CCAGAGAGGCCCACGCACGA[G/C]TCCGGAAGTGAGGGAGGGAG	GAGGCCCACGCACGA	CCTCCCTCACTTCCGGA	1941	22
WIAF-3881	A/G	A	G	AGCTGTAACAGGTTTGGAGA[A/G]GGCAGCTGTGGATTCTCTTT	CAGCTGTAACAGGTTTGGAGA	AAGAGAATCCACAGCTGCC	1942	9
WIAF-3882	A/G	A	G	GCGCTTTATGCCATGTAGAG[A/G]AGACACTGGGTAACCACAGC	GCGCTTTATGCCATGTAGAG	GCTGTGGTTACCCAGTGTCT	1943	11
WIAF-3885	G/A	A	G	CAAATGCGACGAACCTCTGA[G/A]CATCCTGGTGAGGAATAACA	AATGCGACGAACCTCTGA	TTGTTATTCCTCACCAGGATG	1946	1
WIAF-3887	G/A	A	G	AATTTATGTTGCTGATAATA[G/A]CTCAGAAGTACTCTTGGATG	CAAATTTGATAAATTTATGTTGCTGA	TCTATTCATCCAAGAGTACTTCTGAG	1948	14
WIAF-3889	C/T	C	T	AGCCACCAGCACTCATGGGG[C/T]AGTCCCTGTCCAGGCTGCCC	CACCAGCACTCATGGGG	CAGCCTGGACAGGGACT	1950	17
WIAF-3890	A/G	A	G	TCCTTATACAACGCACAAAT[A/G]ACATTAAGCATTATCCCTGA	CATTCCTTATACAACGCACAAA	CTTTTCAGGGATAATGCTTAATG	1951	15
WIAF-3891	A/G	A	G	GTTAGTATACAGATATTACT[A/G]TACTGGTACCCATCTTGTAA	CAATACAAAGCCCTAGTTAGTATACAGAT	TCCTTACAAGATGGGTACCAGT	1952	1
WIAF-3894	C/T	C	T	ACTGATGGTGACGATACTGA[C/T]GATAGTAATAACACTTACTG	CACTGATGGTGACGATACTGA	GGGCACATAGTGAGTATTCAGTAAG	1955	2
WIAF-3896	T/C	C	T	GCTGCTGGGAAAGACAGCAG[T/C]CTTGGCGCGTCCCACAATCC	TGCTGGGAAAGACAGCAG	GTGGGACGCGCCAAG	1957	19
WIAF-3899	C/G	C	G	ATACACGAGGATGAAGTAGA[C/G]GATATGGAAGAAAATTAAGT	TGAAGATACACGAGGATGAAGTAG	AACTGGATCACACTTAATTTTCTTC	1960	10
WIAF-39	T/C	C	T	AATCTGCTGAGAGCAGAGCT[T/C]AAGATCCACAATTGCAAAGG	AATCTGCTGAGAGCAGAGCT	CCTTTGCAATTGTGGATCTT	2451	17
WIAF-3901	G/C	C	G	AAGCATGTGGGGCGTGCTGA[G/C]TGCCCTGGGAAGGTCTGGAG	TGTGGGGCGTGCTGA	AGACCTTCCCAGGGCA	1962	10
WIAF-3902	G/C	C	G	TAGAAACCTCCTGAGTAACA[G/C]AGTTTATGGTTTTACAAAAG	GCATTATAGAAACCTCCTGAGTAACA	GGAGTTCTCTTTTGTAAAACCATAAA	1963	3
WIAF-3906	G/C	C	G	CCCTACGGACGGCGCAATGA[G/C]ATCTGGCTGTTGAAGACATG	CGGACGGCGCAATGA	TCATGTCTTCAACAGCCAGA	1967	12
WIAF-3907	T/C	C	T	AGTGATTATCAAGCAGTCAG[T/C]TTATCTCTTTTCAAACATTT	TGAGAGTGATTATCAAGCAGTCAG	GTCTTTATGAAATGTTTGAAAAGAGA	1968	15
WIAF-3908	G/A	A	G	CTTTTAAATTTTAACCTAAC[G/A]GAAAGCCATTAAAACATTTT	ACAAAGGACTTTTAAATTTTAACCTAAC	GCTTAAAAATGTTTTAATGGCTTTC	1969	10
WIAF-3909	C/T	C	T	AACATTCTTCTGAGAGGACT[C/T]TTCCATGCTCACAGGCACAT	CCAACATTCTTCTGAGAGGACT	TGCCTGTGAGCATGGAA	1970	3
WIAF-3911	G/T	G	T	TACAAGCCAGGAGGGCCAAG[G/T]CGCTGACGACTGCTGCTGTG	GCCAGGAGGGCCAAG	GCAGCAGTCGTCAGCG	1972	12
WIAF-3915	C/T	C	T	GTTCCCCACCCCGTCTTCAA[C/T]AGTAACTGGGGCAGTGGCTT	CCCACCCCGTCTTCAA	GCCACTGCCCCAGTTACT	1976	12
WIAF-3918	G/A	A	G	GCAACCCTGGGCCCCACACC[G/A]TGTGGACAGGCAGACACCAG	CCCTGGGCCCCACAC	GGTGTCTGCCTGTCCACA	1979	22
WIAF-3921	T/C	C	T	TTGGTCAGAACATTGTCCTA[T/C]CTACTTTCTGCATTTGTTTT	CATTTGGTCAGAACATTGTCC	GGCTGGAAAACAAATGCAG	1982	1
WIAF-3923	T/C	C	T	TCTGTCAGCTAACCTTTGCT[T/C]TCATAACCTGCATAAACTTG	GTTCTGTCAGCTAACCTTTGCT	TGCAAGTTTATGCAGGTTATGA	1984	4
WIAF-3924	G/A	A	G	AGCTTGATTTAGTTCATCTC[G/A]TTGTCTCTGGGACAGCACCA	GCTTTAGCTTGATTTAGTTCATCTC	GGTGCTGTCCCAGAGACAA	1985	2
WIAF-3927	C/T	C	T	GAAAATCAGAGATGGTGCCA[C/T]CACCAGACAACACAAGTGCT	GAAAATCAGAGATGGTGCCA	AGCACTTGTGTTGTCTGGTG	1988	20
WIAF-3928	A/G	A	G	CCTGGGGCCAGGTTTGAGCC[A/G]CAGGGAAGTGGGCAGGGGTG	GGGCCAGGTTTGAGCC	CCTGCCCACTTCCCTG	1989	9
WIAF-393	G/A	A	G	TAAAGCAAGGGGACCTTGGC[G/A]CTCTCAGCTTTCCCTGCCAC	GCAAGGGGACCTTGGC	GGCAGGGAAAGCTGAGAG	2704	3
WIAF-3937	C/T	C	T	GGCCTGTCTGCGAGGCCCCA[C/T]GCAATCGTCGCAAACAAGAT	GTCTGCGAGGCCCCA	TTGTTTGCGACGATTGC	1998	20
WIAF-3938	A/C	A	C	GATGTATCTTCTGATTTTCC[A/C]ATTTCCCGAAGTTGCCCACA	CCAGGATGTATCTTCTGATTTTCC	GGGCAACTTCGGGAAAT	1999	9
WIAF-3939	A/G	A	G	CCCTCTGTTTGGGACGGAAC[A/G]GCTGCAGCGGGTGAAGGAAG	CCTCTGTTTGGGACGGAA	TTCACCCGCTGCAGC	2000	20
WIAF-394	A/C	A	C	CCCAGCCTGGCCAGACAAGA[A/C]GACTGTCAGGAAGGGTCGGA	AGCCTGGCCAGACAAGA	CCGACCCTTCCTGACAGTC	2705	3
WIAF-3941	T/C	C	T	GGGCCTGACGTCACTAACGG[T/C]AACTGACAATCTTGGAATGG	GCCTGACGTCACTAACGG	TCCATTCCAAGATTGTCAGTT	2002	20
WIAF-395	T/C	C	T	TCTGAGTCCTGAGACTTTTC[T/C]GCGTGATGGCTATGCCTTGC	TTTATGGTGTAGTTCTGAGTCCTG	AGGCATAGCCATCACGC	2706	3

续表

WIAF 命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi 编号	NwLib chrom
WIAF-402	A/G	A	G	CCAACTTACCAAACCTCTGT[A/G]GCTTAGCCTCGCCTACCGTA	CCCAACTTACCAAACCTCTG	CGGTAGGCGAGGCTAAGC	2709	14
WIAF-4024	A/G	A	G	CCTCCTCCGCATGCCAACGC[G/A]GTTCATGTACAAGGCCCCTC	GATGTGTTGCCCCTC	AGAGGGGCCTTGTACA	1002	4
WIAF-4027	A/C	A	C	CTGCAACCCACTCTGAGCCT[C/A]TCTCTCCTCCTATTTTACTT	TCTGCAACCCACTCT	CAAGTAAAATAGGAGGAGAG	1004	1
WIAF-4028	A/T	A	T	TTGAAAAGTCAGGTACAAAT[A/T]TCCTCAAGTGGCATAAAAAT	TTTGAAAAGTCAGGTACAA	CATTTTTATGCCACTTG	1005	10
WIAF-4029	C/T	C	T	CAGCTCCTTGACCTATGAGC[C/T]GGGGCCTGACTAGGAAAAGT	TCAGCTCCTTGACCTATG	AACTTTTCCTAGTCAGGCC	1006	1
WIAF-4035	C/T	C	T	AGATTCACGTGGAGAGATGT[C/T]TTGGCCAGGGCGGGCAGATG	AGATTCACGTGGAGAG	GTGGGCTCACATCTG	837	8
WIAF-4045	A/T	A	T	AATGGGCAGTGGCACTAGAA[A/T]AATCTTGAGCACAGTGAATG	ATCAATGGGCAGTGG	CATTCACTGTGCTCAAGAT	1010	12
WIAF-4050	C/G	C	G	CCTTTCATAAAAACCAAACT[C/G]TAGCAAGATGCAAATGCATG	CCCTTTCATAAAAACCAAAC	CCATGCATTTGCATCT	1014	1
WIAF-4055	A/G	A	G	TTACCATAACATATTCATGG[A/G]TTCTGGGATAAGGGGTAGAC	TTGTCATTACCATAACATATTCATGG	TGTCTACCCCTTATCCCAGA	798	10
WIAF-4066	A/G	A	G	CGCTCTAGCCTCTAATTCCC[G/A]CTCTAGAACAGCTGGCCCTG	CGCTCTAGCCTCTAATT	GTACTGACGACCAGGG	1017	1
WIAF-4070	C/T	C	T	AGAACTAATACCTGCTGTTC[C/T]TCTGCCTCCACCAGCCCTAT	CAGAACTAATACCTGCTGT	GATAGGGCTGGTGGAG	840	10
WIAF-408	C/T	C	T	TTTAAAGCGAAGTTGAAACA[C/T]GAAGACGATAGTTAACGTCT	TTACTTTAAAGCGAAGTTGAAACA	TTACCAGACGTTAACTATCGTCTTC	2140	17
WIAF-4080	A/C	A	C	GGGCAGCACCCTATAAATCA[A/C]CACCTTAATATGCTGCAACA	GGGCAGCACCCTATAA	TTGTTGCAGCATATTAAGGT	1028	18
WIAF-4082	A/G	A	G	CACCTCTGCTGCCTTGGTTC[A/G]AGCCCTCATCTCTTTTACAG	TCACCTCTGCTGCCT	CCTGTAAAAGAGATGAGG	19300	14
WIAF-409	G/T	G	T	ATACGGTGTGCGAGGCAACA[G/T]GGAGAGGTACGGGAATAGTT	GGTGTGCGAGGCAACA	GGAAGTAGAACTATTCCCGTACCTC	2141	17
WIAF-4095	C/T	C	T	ACGTTCTTCCACCAGAATTT[T/C]CGAGCCACCTTCTGCGATGT	ACGTTCTTCCACCAGAA	GACATCGCAGAAGGTG	1033	2
WIAF-4098	C/T	C	T	GTTTCTTCATAAGATGGAAA[C/T]GCTATACCTTACCTACCTCG	TTCAGTTTCTTCATAAGATGGAAA	TCAGACTTTTACGAGGTAGGTAAGG	19449	3
WIAF-4117	A/G	A	G	ATCTTTCCCTTTTCATCTCC[G/A]TTTGTGTGTTTGGCCAAATA	CATCTTTCCCTTTTCA	TTATTTGGCCAAACA	843	7
WIAF-413	A/G	A	G	TGCCCTTCACATAACAGAAT[A/G]ACTTGCCATCTGCCTGCACC	TGTGCCCTTCACATAACAGAA	GCAGGCAGATGGCAAGT	2715	1
WIAF-4131	C/T	C	T	TGATAGAATTGCTAAATTGT[T/C]GTGAAATAGGTTAGAATTTT	CGGTGATAGAATTGCT	AGAAAAATTCTAACCTATTTCA	1042	9
WIAF-4138	T/G	G	T	AGTTAGGGTCTATGTATTTA[G/T]GATGTCTGCACCTTCTGCAG	GCAAGTTAGGGTCTATGTA	CTGCAGAAGGTGCAG	1043	13
WIAF-4139	C/T	C	T	CAGTGAAGGAAGGGATGGCA[C/T]AGCGTTATGTGTAAAAAACA	CAGTGAAGGAAGGGATG	TGTTTTTTACACATAACGC	1044	1
WIAF-414	C/T	C	T	GGAAGACACCGTGCAAATGC[C/T]AAAGTGCACTGAGGAGAGGG	AAGACACCGTGCAAATGC	CCCTCTCCTCAGTGCACTTT	19420	8
WIAF-4142	C/G	C	G	AGAGACAGCGATTGGCTAAC[C/G]CATGGCAGTAGTGGGCCCCA	AAAGAGACAGCGATTGGCTAA	GGGCCCACTACTGCCAT	806	4
WIAF-4147	C/T	C	T	TCAGAGGATACAGAGGCAAA[C/T]GTTGGTTATAGCAGTCAACG	TGGGTAAATGTTATAGAAACTTCAGA	TGTCGTTGACTGCTATAACCAA	846	18
WIAF-4150	T/G	G	T	CATTCACACATCCCCTTTCT[T/G]AGATGGTATTGGAGAAGTAG	CATTCACACATCCCCTT	TCTACTTCTCCAATACCATC	19187	8
WIAF-4164	T/G	G	T	CCTCTGTGCTCTTCTCTTCC[G/T]GTCGGATCTCTCTCCTCTCT	CCTCTGTGCTCTTCTC	AGAGAGGAGAGAGATCCG	1055	17
WIAF-4186	C/T	C	T	CTCCCACCTTTAATCTGTTT[C/T]ATACCTTGCTTATTAAATGA	GCTCCCACCTTTAATCTG	CGCTCATTTAATAAGCA	1062	19
WIAF-4187	T/G	G	T	GTTTTAGATAAAGGGCACTT[G/T]GCAGGAGTGTTTAGGATGAA	CATGTGTTTTAGATAAAGGGCACTT	TCTCTTCATCCTAAACACTCCTG	19386	14
WIAF-4189	C/T	C	T	TTTATCGCTATAGATTTGCC[C/T]ACTTGACATATCATACACAT	ACTCTTTATCGCTATAGATTTGCC	GGCTCCATGTGTATGATATGTCA	19450	
WIAF-4216	A/G	A	G	AGCTACAAAACCTACCCTCA[A/G]TGAGCATGGTACTTGGCCTT	GGAGCTACAAAACCTACCCTCA	CAAAGGCCAAGTACCATGCT	19606	14
WIAF-422	A/C	A	C	CCTCCCCTGCCCCCATCCCA[A/C]ATGATCTTGAGATTTCTTTT	CCTGCCCCCATCCCA	CTTCTTTAAAAGAAATCTCAAGATCAT	2718	17
WIAF-4222	A/C	A	C	ACAATTACCCTGTCAAGAGG[C/A]GAGTGCAGCTCAGGTGGATT	AACAATTACCCTGTCAAG	AATCCACCTGAGCTG	1074	4
WIAF-423	T/A	A	T	GAGGTGGGAGACAAGGAACC[T/A]CCGAAGAGGAAGCAAGAAAG	GGTGGGAGACAAGGAACC	TTTCTTGCTTCCTCTTCGG	2719	17
WIAF-4237	C/G	C	G	ACACCTTCCCCACTCTCTTA[C/G]GGTACAGAAAGGAGATGCAT	GTACACCTTCCCCACTCTCTT	TGCTGTTCATGCATCTCCTT	682	1
WIAF-4240	A/G	A	G	AAGCAGATCTTGAGGGTTAT[A/G]GTTAAGCCTGATAACAGCCT	TGTCAAGCAGATCTTGAGGG	AAAGAGGCTGTTATCAGGCTTAAC	19363	11
WIAF-4242	C/G	C	G	AGTCACCATGCCCAGCCTAG[C/G]ATGAGTTTAGTAAGATTTGG	AGTCACCATGCCCAGCCTA	CTCCCCAGCATAACCAAATC	19369	22
WIAF-4245	A/G	A	G	ATATTTTCTTGACATGAGGT[A/G]GCTTTTTAGCAGCATTTCGG	AAATCCATATTTTCTTGACATGAGG	CCGAAATGCTGCTAAAAAGC	19156	10
WIAF-425	C/G	C	G	ACACTGAAGACTCACCAGAA[C/G]AGGGTGGGGTGGGGAATACT	CCACACTGAAGACTCACCAGA	TCCCCACCCCACCCT	2721	15
WIAF-4251	A/G	A	G	GCCGACATCAGCATTGTCTC[A/G]TGTACAGCTCCCTTCCCTGC	TCAGACACTGCCGACATCA	TCTTGTCCAGTTGTCTCCCA	19415	3
WIAF-4253	C/T	C	T	GGGGTGAGCACACAGCAAAA[C/T]GGGGTGGGACGTGCAGAGAG	GTACACATCCTGGGGGTGAG	GCTTCCTTTGCCTTTACCCT	19382	
WIAF-4256	C/G	C	G	TGCACTGTTACTCCCCAGAC[C/G]GAGAGCTTACATACCATATA	TGCACTGTTACTCCCCAGAC	CGATCCTACACATTCCTTCTGA	19383	19
WIAF-4257	A/G	A	G	TTATTTCATATTCCCACCAC[A/G]ATAACGACTCCTTTAATTTA	TTCTACTTTATTTCATATTCCCACCA	CCCTTTCAGGAACCCTGTAT	19384	6
WIAF-4259	A/C	A	C	CTGAATCACACAACATGGAC[A/C]AATCTCAAATCATTATGCTG	CCACTGAATCACACAACATGG	TCCATCAGCATAATGATTTGAGA	19607	14
WIAF-432	A/T	A	T	GAGGGTTGAGGTGTAGATAT[A/T]CTTCCTCTCTTCTCGTGGAG	CTAGGAGGGTTGAGGTGTAGATAT	GCTCCACGAGAAGAGAGGAA	2727	11
WIAF-433	A/G	A	G	GCTTGGACCTTGGATAGGTC[A/G]ACCGGCTGAAGGTTGGACAG	GCTTGGACCTTGGATAGGTC	CCAACCTTCAGCCGGT	2145	9
WIAF-436	T/G	G	T	CAGATCTCTGGTTTATTTAA[T/G]ATCAACATTCACCACAGCTG	TGAAGATCAGATCTCTGGTTTATTT	CAGCTGTGGTGAATGTTGAT	2730	14
WIAF-440	C/T	C	T	TTGAAAAAGTCTTTCAGCAC[C/T]TTCGTGGATCCGAATTTTAG	CAGGACAAACTTTTGAAAAAGTC	CTAAAATTCGGATCCACGAA	2148	X
WIAF-444	T/C	C	T	AGCCTCCCTTTAGACACCTC[T/C]ACAGGTACAGCCGACCATGC	AGCCTCCCTTTAGACACCTC	CATGGTCGGCTGTACCTGT	2734	15
WIAF-4500	G/C	C	G	TGCAAAAAGGGTCAGTATGG[G/C]TTAGGGAAAACATTCCATCC	TGCAAAAAGGGTCAGTATGG	AGGATGGAATGTTTTCCCTAA	19421	2
WIAF-4501	C/T	C	T	TCTTCCTGCTTTATTTTGCC[C/T]ACAGTTTAGGTAAATAAAAG	TCTTCCTGCTTTATTTTGCC	TCTTGGGCATCTTTTATTTACC	19403	1
WIAF-4502	G/T	G	T	CACTTAAGCCATTACCAATA[G/T]TGAGATGTAATGGAGAATTT	GCTCTTTACACTTAAGCCATTACCA	CCACATTAAATTCTCCATTACATCTC	19387	2

续表

WIAF 命名	等位基因	等位基因 A	B	芯片中的位置	上游引物	下游引物	ncbi 编号	NwLib chrom
WIAF-4503	C/T	C	T	CAAATGGTGCCATAGTATAA[C/T]AATAGCTATAACAATTATGG	CCAAATGGTGCCATAGTAT	CAATCCATAATTGTTATAGCT	19432	7
WIAF-4504	A/G	A	G	CCATAATTGTTATAGCTATT[A/G]TTATACTATGGCACCATTTG	AAGCAATCCATAATTGTTATAGCTATT	TCCCAAATGGTGCCATAGTA	19389	7
WIAF-4505	G/A	A	G	ATGCTTCCTCCTTTGCTCTT[G/A]GGAAGACCAGCTTTGCAGTG	CATGCTTCCTCCTTT	CACTGCAAAGCTGGTC	19404	11
WIAF-4506	A/G	A	G	TATTCCAGATGTGTTTGGTC[A/G]TGCGTATGGCAGTGAGCAGG	TATTCCAGATGTGTTTGGT	ATACCTGCTCACTGCC	19608	7
WIAF-4507	T/C	C	T	TATTTGTTCTGCTCATAATT[T/C]CCAATATGTACCAGACCTTC	TTAGTTGAGTATTTGTTCTGCTCATAA	GGGAAGGTCTGGTACATATTGG	19422	1
WIAF-4508	C/T	C	T	CGGAGAAGCATTGTTTTCTT[C/T]GGCCAAGAAGGTATCTACC	ACGGAGAAGCATTGTTTT	GGTAGATACCTTCTTGGC	19390	9
WIAF-4509	A/G	A	G	TGGGTAGGTCCAGTAATCCA[A/G]GGATTCAGGAAGGAGGCCAC	CTGGGTAGGTCCAGTA	CTTCGATCCTCGTGG	19391	8
WIAF-4510	C/A	A	C	AGGGAGACAGAATGCTGATT[C/A]TCTGGTGGAGAACCAGAACT	AGGGAGACAGAATGC	AAGTTCTGGTTCTCCACC	19344	19
WIAF-4511	A/T	A	T	TGTGGAGCTAGAATCCCCAG[A/T]GGCCAGCTGAGTGCTGGTCG	TGTGGAGCTAGAATCCCCAG	CCAGATGTGCTGGCTGACT	19377	9
WIAF-4512	A/G	A	G	GAAATAGGGACTGCAAGGGC[A/G]GGCTGTATTGAGAAGGATGG	TGAAATAGGGACTGCAAGGG	GGGCCATCCTTCTCAATACA	19364	17
WIAF-4513	C/T	C	T	TAACACATTTTTCTGCAAAT[C/T]ACCTCTTTCATTTAACAGCC	GTAACACATTTTTCTGCA	GGGCTGTTAAATGAAA	19392	15
WIAF-4514	G/A	A	G	CAGAAATCCAGTTATTTTCC[G/A]CCCTCAAAATGACAGCCATG	CCAGAAATCCAGTTATTTT	CATGGCTGTCATTTTGA	19393	21
WIAF-4515	A/C	A	C	ACATGGCTTCGAGAGAGTTG[A/C]ACAGATTCCTGGAAGACAGC	AACATGGCTTCGAGAGA	GCTGTCTTCCAGGAATCT	19394	15
WIAF-4516	T/C	C	T	CTTGCCTCCTTCTGGTGCTC[T/C]TACTCCAAGTCTATTTCATT	TGCCTCCTTCTGGTGCTC	GGAAAAATGAAATAGACTTGGAGT	19451	8
WIAF-4517	A/G	A	G	TGTTTTCTATTACAGGACAC[A/G]TGGATGCTGGAATCACCCAG	TTGTTTTCTATTACAGGACA	CTGGGTGATTCCAGC	19345	7
WIAF-4518	C/T	C	T	CCAGTGACTCTGAGATGTCA[C/T]CAGACTGAGAACCACCGTTA	ACCAGTGACTCTGAGATGTC	TAACGGTGGTTCTCA	19346	7
WIAF-4519	C/A	A	C	GAAAATATGATTAGAAAAAG[C/A]GTGGTACATCTTAAAATATA	GATTTCAGAAAATATGATTAGAA	GAAAGTTATATTTTAAGATGTACCA	19405	10
WIAF-452	G/A	A	G	TTTGCTTTTTAAATAGAACA[G/A]CTTTGATTTTTAGTATATGA	CTTTTCATTTTTGCTTTTTAAATAGAAC	TGATGATGTCATATACTAAAAATCAAAG	2739	4
WIAF-4520	G/C	C	G	ATGAAATAAAAGTAAGACCA[G/C]ATAAAAATACCTATGATGCT	ATAATGGTATGAAATAAAAGTAAGACCA	AACAATGAGCATCATAGGTATTTTT	19336	13
WIAF-4521	A/G	A	G	AATCTGAATATTCAGCACAT[A/G]CAAGTGTGACAACCACTTGT	GGAGAATCTGAATATTCAGCACA	TACTAAACAAGTGGTTGTCACACTT	19452	7
WIAF-4522	A/G	A	G	TGCCAGGAGAGTAATACTAT[A/G]CATTACAGAGAAAAGACAAC	TTCATTGCCAGGAGAGTAATACT	GTTTCAGTTGTCTTTTCTCTGTAATG	19453	4
WIAF-4523	C/G	C	G	CAGATCATTGGTTAGGCTCA[C/G]CTTCCTGTAATTGCTTCTGT	CAGATCATTGGTTAGGCTC	AACAGAAGCAATTACAGGAA	19406	9
WIAF-4524	C/T	C	T	ACCAGATTATTAGGAGACAG[C/T]CATGCAAAGGAATTTCTAGC	CCCCTACCAGATTATTAGGAGACA	TGCTAGAAATTCCTTTGCATG	19454	8
WIAF-4525	T/A	A	T	TAGCATGAGAGACAATTTAC[T/A]TTTTTAAGAACTTTGTGATC	ACTCTAATTAGCATGAGAGACAATTTAC	GGCTGATCACAAAGTTCTTAAAAA	19455	11
WIAF-4526	T/A	A	T	GAGTACTGACCCGGGAGATG[T/A]CTTCGAAAAGGAAAACCAGT	AGTACTGACCCGGGAGATG	CACTGGTTTTCCTTTTCGAAG	19456	12
WIAF-4527	C/T	C	T	GTGTTTTCAGCAAAAACTTG[C/T]ATAGATCATTCAGAGAGGGC	CTGTGTTTTCAGCAAAAACTTG	CGTAGCCCTCTCTGAATGATC	19457	4
WIAF-4528	G/A	A	G	GAGGAGTGTTTTTTCTATCT[G/A]CAGCCAAACAGAAATACTGT	ATACTGGGGCCATTGTAGG	TGTACTACAGTATTTCTGTTTGGCTG	19458	4
WIAF-4529	T/C	C	T	GTTTACGGATATTTATATTC[T/C]AGTCTCACTCTCTTGTCCCA	AGATGCTGTTTACGGATATTTATATTC	GGTGGGACAAGAGAGTGAGACT	19407	1
WIAF-4530	C/T	C	T	TTACATGGTGAATAACTTTC[C/T]ATAGGACCCACAAGGCCTGC	TGCTAGATTACATGGTGAATAACTTTC	CAGGCCTTGTGGGTCCTAT	19459	11
WIAF-4531	T/G	G	T	TTTCATCCTTTCAACATCTT[T/G]TGTCACATTTTAGGTGATGC	TTTTTGGTTAGTGTATGTCTACTATACCTT	TTACAAGAGCATCACCTAAAATGT	19430	14
WIAF-4532	C/T	C	T	TTCCAGAGAAAATGGGATTG[C/T]CTGCAGCGCCCTTCCTCACC	TTCCAGAGAAAATGGGAT	TTTGGTGAGGAAGGG	19290	6
WIAF-4533	A/G	A	G	TTATGTACAAGAAAATGGAA[A/G]TGATTTTAGATCCTCCCCCA	TGCTTTATGTACAAGAAAATGGAA	CTGGGGGAGGATCTAAAATC	19425	5
WIAF-4535	G/A	A	G	TAGTTGTGCTCCAGGACCTA[G/A]GCGTGCTCACTCTACCTTGT	CTAGTTGTGCTCCAGGAC	ACAAGGTAGAGTGAGCA	19460	2
WIAF-4536	C/T	C	T	ACATCCACTGCTTTCAATAA[C/T]TCACTCCACTTATGTCTTAC	CCTCACATCCACTGCTTTCA	TGTAATTTTGTAAGACATAAGTGGAGTG	19395	8
WIAF-4537	G/T	G	T	AGGATCTTTCAGGTAGCACT[G/T]GAGAATCTGAATATTCAGCA	TCAGGATCTTTCAGGTAGCACT	TTGTATGTGCTGAATATTCAGATTC	19424	7
WIAF-4538	C/T	C	T	AGGAAATGACTGGATAATGA[C/T]AGAAATGAATAGAGCCCCAT	TCAGGAAATGACTGGATAATGA	AAAATGGGGCTCTATTCATTTC	19426	2
WIAF-4539	C/T	C	T	ATGTCCATTACCTCTACTTA[C/T]GACAAGCAAGAACAACAACA	GAGGTCCATGTCCATTACCTCT	TTCTGTTGTTGTTCTTGCTTGT	19408	7
WIAF-4540	C/T	C	T	CGTTAACATGCATGTCTGTT[C/T]AACAAGTGTTTGTGGTGTCA	AACGTTAACATGCATGTCTGTT	TGATGACACCACAAACACTTG	19429	9
WIAF-4541	A/G	A	G	GCCTATGTTACTGGTATGCT[A/G]TGGTATTGGATTGGGATGGA	CCCAAGCCTATGTTACT	CAAGTAATCCATCCCAATC	19388	2
WIAF-4542	A/G	A	G	GTAAGATGTGAACCTATACA[A/G]TGGCAAGGAAAGTAGAAATG	TTTGGAAGGTAAGATGTGAACC	TCCATTTCTACTTTCCTTGCC	19427	14
WIAF-4543	A/C	A	C	CACTGCCGCCATATCTCCTC[A/C]CCAACACCTCTGTTTTCTGA	TGCCGCCATATCTCCTC	CTGTCAGAAAACAGAGGTGTTG	19428	1
WIAF-4544	A/G	A	G	CAGGCCCACTGTTCCTATGC[A/G]CACTGGCTTTGTAGGCATTC	AGGCCCACTGTTCCTATGC	TGAATGCCTACAAAGCCAGT	19409	9
WIAF-4545	C/T	C	T	GGAAAGGGGTCAGCTAAAGG[C/T]AACTGTTCTATAAGGATGGG	GGAAAGGGGTCAGCTAAAGG	GGATACCTACCCATCCTTATAGAACA	19410	12
WIAF-4546	A/C	A	C	CCTGGTGCTGATTGCCTAGC[A/C]GGAGAGTTGAGTGCCACAGG	CTGGTGCTGATTGCCTAGC	ACCTGTGGCACTCAACTCTC	19411	21
WIAF-4547	G/A	A	G	AGCAAAAGGTTGTTCCTAAA[G/A]TAAGGGCAGAGTCACACGGG	CAGCAAAAGGTTGTTCC	TGTATCAGCTGCCCC	19412	15
WIAF-4548	G/T	G	T	AAAGGTAGCTTTTCCCAAAC[G/T]GATCTTTTCATTTAGGTGAG	AAAGGTAGCTTTTCCCAAA	CCTCACCTAAATGAAAAG	19413	12
WIAF-4549	A/G	A	G	TACCTTCTGGGGCATACAAC[A/G]TGGCAGCAGGGCCTCGGGAA	GTACCTTCTGGGGCATAC	CTACCCCTCTTCCCG	19396	2
WIAF-455	T/C	C	T	GAATCCCCTATCCCATTCTG[T/C]GTATGAGTCCCATTTGCCTT	GAATCCCCTATCCCATTCTG	CAAGGCAAATGGGACTCATA	2741	1
WIAF-4550	C/G	C	G	GAGCACCCACCTACCTCATG[C/G]AACTGTTGAGAGAAATAAAT	GAGCACCCACCTACCTCATG	CAATGTCTCATTTATTTCTCTCAACA	19357	X
WIAF-4551	C/T	C	T	TGGACAACAAAAAGGAATA[C/T]GATCCAGCCTCATGCAACAT	TCATTTAATCATGGACAACAAAAA	TGTTGCATGAGGCTGGAT	19183	15
WIAF-4552	T/C	C	T	CTCATCATCTGCGCAAGAGA[T/C]CAAAGAGGTCAGCTTCTGTT	CTCATCATCTGCGCA	CAACAGAAGCTGACCTCTT	19207	22
WIAF-4553	T/C	C	T	TAGAGAGTAGCAGTATTTCA[T/C]TGGTACCTGGTACTGATTAT	ATCTGTAGAGAGTAGCAGTATTT	CATAATCAGTACCAGGTACC	19347	17
WIAF-4554	A/G	A	G	TAGAGAAAATGTTTTTAAAG[A/G]AGCCAGCTCAAGCAATATTA	CGTAATAACATTAGAGAAAATG	TTAATATTGCTTGAGCTGG	19348	17
WIAF-4555	A/G	A	G	TTTTAGCAAAAGCGTCCAGA[A/G]AGGAGAGCTTAGCAGGAGTC	TTTTTAGCAAAAGCGT	GGACTCCTGCTAAGC	19349	17
WIAF-4556	A/G	A	G	TTATTTTGGAAGGATGCCCC[A/G]GTATGTCTCCGACCACTTGC	TTATTTTGGAAGGATGCC	AGCAAGTGGTCGGAG	19350	
WIAF-4557	C/T	C	T	AGCCAGAGTCTGCACAGTCA[C/T]AGGGCAAGCAGAAAATTCTT	GGTTGCTCGAGCCAGAGTC	GAAAGAATTTTCTGCTTGCCC	19461	11

续表

WIAF 命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi 编号	NwLib chrom
WIAF-4558	A/T	A	T	GCTTCCTTTGGCGAAAGGAT[A/T]AAGAAGTGAGTGACGGTGAC	GATGAAGTGGCTTCCTTTGG	CACAGGTCACCGTCACTCAC	19370	5
WIAF-4559	A/G	A	G	AGATGGAGACGAAGGCCAGC[A/G]TGGCTCTCACAGGGCAGCGC	TCAACCAGAGATGGAGACGA	AGCGCTGCCCTGTGAGAG	19282	
WIAF-456	G/C	C	G	GACTCCTGGGACAAAAGGAC[G/C]AAAAACACTCTTCTGTGGAT	GACTCCTGGGACAAAAGGAC	CAGATATCCACAGAAGAGTGTTTTT	19161	3
WIAF-4560	G/T	G	T	AGGGCTCAGAAGATGTCGGG[G/T]ACGATATAATCAGTGTAGGC	AGAGGGCTCAGAAGATGTCG	CATCGATGGTGCCTACACTG	19367	3
WIAF-4561	C/T	C	T	ACAGACACAGCATCACACCA[C/T]AGGGCCCACGGGAGGGTCGG	ACAGACACAGCATCACACCA	GATTAGAGCTGCCTTTCCCA	19416	1
WIAF-4562	G/A	A	G	TGGGTGCTGTTCTCTGCCTC[G/A]GGAGCTCTCTGTCAATTGCA	TTGGGTGCTGTTCTCT	TGCAATTGACAGAGAGC	19351	2
WIAF-4563	C/T	C	T	GGTCTAAGTTGCTGATTACC[C/T]GGATTTTTCTGACGATCTTT	TGGTCTAAGTTGCTGATT	GAAAGATCGTCAGAAAAA	19352	7
WIAF-4564	G/A	A	G	ATCCACACTGCACACTGCCT[G/A]GCCCAAAACGTCTTATTGTG	TATCCACACTGCACACTG	CACAATAAGACGTTTTGG	19397	14
WIAF-4565	C/T	C	T	TTTAAGAAGCTTGATTTGGA[C/T]AATTCTGGTTCTTTGAGTGT	TTTAAGAAGCTTGATTTGG	CACACTCAAAGAACCAGAA	19398	2
WIAF-4566	G/A	A	G	ATGCAGAAAGAGGGGAAAAA[G/A]GAGCGAGCTGTGGTGGACAA	AATGCAGAAAGAGGG	CTTGTCCACCACAGCT	19353	1
WIAF-4567	G/C	C	G	TAGGATAAGTGTTATCCCAG[G/C]TTCATAGGTATGGAGTCTCA	AATGAGTAGGATAAGTGTTATCCCAG	CATCTATGAGACTCCATACCTATGAA	19243	5
WIAF-4568	C/A	A	C	AATTTCATTTATCATCTGGA[C/A]AGCCCCTTCTTATAACGTAC	GGGGTGGTTTATTAATTTCATTT	GATGTACGTTATAAGAAGGGGCT	19462	1
WIAF-4569	G/C	C	G	TACCGCTGCTATTCCCAGAT[G/C]AAGATTTGGTGGAAGGAGAC	CCGCTGCTATTCCCAGAT	GGTCTCCTTCCACCAAATCTT	19419	9
WIAF-457	A/G	A	G	ACCTCTTGTTGGATAAAAGG[A/G]CATTGTTTTTCATTAGCTTG	TTAATACCTCTTGTTGGATAAAAGG	GAAGACAAGCTAATGAAAAACAATG	2742	8
WIAF-4570	G/A	A	G	GAAAATACAAGAACCTGCCC[G/A]AAATAAGGCAACTTTTGGAA	AGAAAATACAAGAACCTGCCC	TCTTCCAAAAGTTGCCTTATTT	19463	3
WIAF-4571	G/A	A	G	GAGTGAGGGCATAATAGGCT[G/A]TGGTGTTAGAAGGCAGGAAT	AGAGTGAGGGCATAATAGGCT	AAATAACCACATTCCTGCCTT	19464	2
WIAF-4572	A/G	A	G	AAAATTCTATCCACATCCAA[A/G]GTGTCTCTGAGAAGTCGAGG	GCAAAATTCTATCCACATCCAA	GCCTCGACTTCTCAGAGACA	19465	6
WIAF-4573	G/A	A	G	GGATAAAGGGACTGAGATGC[G/A]AATGGGGATGGGCAATAACC	TGGATAAAGGGACTGAGATGC	TTATTGCCCATCCCCATT	19466	7
WIAF-4574	T/C	C	T	CAAAATGTAAATCAAAACCA[T/C]ACTTTTCTGCATGTACTAAG	CGTATTCCCAAAATGTAAATCAAA	TGACCTTAGTACATGCAGAAAAGT	19467	6
WIAF-4575	A/C	A	C	TACTTTCAGGAGCAAGCATC[A/C]CATATATGTAGAAGTAAAAT	AATACTTTCAGGAGCAAGCATC	TGGCTTTTTCATTTTACTTCTACA	19468	13
WIAF-4576	C/T	C	T	GACTCATTTCTTCTTTGCAT[C/T]TATTTCTAGGTTATTTGCAG	TGGACTCATTTCTTCTTTGCA	GGGGCTGCAAATAACCTAGA	19469	7
WIAF-4577	C/T	C	T	TCCCCCAACTTCACATTCAG[C/T]AGATATTCTTTCATGGGTTA	GTTCCCCCAACTTCACATT	GCAAAATAACCCATGAAAGAATAT	19470	15
WIAF-4578	C/T	C	T	GAATGTACTTGATGTACTTG[C/T]CCTCACATCCATTTACCACC	CAGAGTGAATGTACTTGATGTACTTG	GGTGGTAAATGGATGTGAGG	19471	3
WIAF-4579	T/C	C	T	CGGACGGAAAGCGGGGCCTG[T/C]GGGTCCTTGTGTCCCTGCCG	GGAAAGCGGGGCCTG	CAGGGACACAAGGACCC	19358	10
WIAF-4580	A/G	A	G	AGATGGGAAAGGTGACCAAA[A/G]ACAATATAGAGGCAGAGGCC	GATGGGAAAGGTGACCAAA	TGGCCTCTGCCTCTATATTG	19472	21
WIAF-4581	A/T	A	T	AAGCAAACACTTACAGAATT[A/T]TGAAGAGGTATCTGTTTAAC	AAGTTAAAAGCAAACACTTACAGAATT	GGAAATGTTAAACAGATACCTCTTCA	19473	7
WIAF-4582	C/T	C	T	GTTTTGAACCCTAGCTGTCA[C/T]GTGCCACCTGCGGGATCTAG	CGTTTTGAACCCTAGCTGTCA	TCCCGCAGGTGGCAC	19474	11
WIAF-4583	C/T	C	T	TAGTCTGTCTCCATTTGCCA[C/T]GTCTTCCTCAGTAGAATAAG	ATAGTCTGTCTCCATTTGCCA	CCCTGTCTTATTCTACTGAGGAAGA	19475	6
WIAF-4584	C/T	C	T	GAGCCTTTAGTGCAAAAACA[C/T]TATGCCATGCGGGAAATAAA	TGAGCCTTTAGTGCAAAAACA	TTTCCCGCATGGCATA	19418	6
WIAF-4585	T/C	C	T	GCATTGCTGGGCTGTGTTCC[T/C]CGGGCTCTTCTGGACCTTGC	TGCTGGGCTGTGTTCC	GGTCCAGAAGAGCCCG	19423	17
WIAF-4586	A/G	A	G	ACATTAAGAACAAGCTTTCC[A/G]CTCCCACTTCCCTCCCACTA	CAGACATTAAGAACAAGCTTTCC	TGGGAGGGAAGTGGGA	19359	2
WIAF-4587	G/A	A	G	CAGGTCAGCCACCACTATGC[G/A]CAGGTTCTCATCATTGAAGC	AACAGGTCAGCCACC	GCTTCAATGATGAGAACCT	19354	22
WIAF-4588	T/C	C	T	CACTCAGGGCCGTCATGCTG[T/C]ACACATCCAGGGGGCGCCCT	ATACACTCAGGGCCGT	TACAAAGGGTAGGGCG	19355	17
WIAF-4589	T/C	C	T	ATGTGGTGGTCTTGTGTAGA[T/C]GGGGGCTTTGGTTTCAGTTG	ATGTGGTGGTCTTGTGTAG	GCAACTGAAACCAAAGC	19356	17
WIAF-459	A/G	A	G	ATAACGTAAGTACCTATGTC[A/G]TGCCATGTAGTTTTTTGGTT	CCCAAAAAATAACGTAAGTACCTATGTC	TTGCAAGTAAATGAACCAAAAA	19609	9
WIAF-4590	C/T	C	T	CCATCAAAATTTCCTTCACA[C/T]TCAATACTGTTGAACAACAA	GCCATCAAAATTTCCTTCACA	TTCAAGTGGGATGAGCAAGA	19379	15
WIAF-4591	C/T	C	T	ATCTCAAGCATCGATGTCAA[C/T]GGGGGCAACCGGAAGACCAT	CATCTCAAGCATCGATG	ATGGTCTTCCGGTTG	19399	
WIAF-4592	A/G	A	G	AAAATACCTCCATTTTGCTT[A/G]TCCTTTTAGTGAAGATGATA	TAAAATACCTCCATTTTGC	GGTATCATCTTCACTAAAAGG	19400	11
WIAF-4593	T/C	C	T	GCATTGTTTCTAGGAGAACC[T/C]GCGTCATACCTTTATCTATA	TGCATTGTTTCTAGGAG	GCTATAGATAAAGGTATGACG	19401	15
WIAF-4594	G/A	A	G	TGGAGACTGCCTGCATGAGG[G/A]TTGGCTTATCGGGGATACAG	GAGACTGCCTGCATGAGG	TGTATCCCCGATAAGCCAA	19476	4
WIAF-4595	G/C	C	G	TAATTCCTTTGTGCAGGAAG[G/C]AGGTTTTGTGGAAATGATGT	GATAATTCCTTTGTGCAGGAAG	CAACATCATTTCCACAAAACCT	19477	15
WIAF-4596	A/C	A	C	TTTTTCCCTTTTAAATGCTA[A/C]ACAAGTGCCACTAATCCACA	GGAGTTTTTCCCTTTTAAATGC	CTGTGGATTAGTGGCACTTGT	19478	2
WIAF-4597	G/A	A	G	TTATCTGGCCAAACTGATAC[G/A]GCATGACCCACAGCCTCAGG	CTTATCTGGCCAAACTGATA	ATACCTGAGGCTGTGG	19360	15
WIAF-463	C/A	A	C	GCTGTGGGTGGTGATTTGGT[C/A]AGCATATCTTAGGTATATAA	TGTGGGTGGTGATTTGGT	CCAGTTAAAAGTTATGGCTTCAAAG	2744	12
WIAF-468	G/T	G	T	TCAGTGTCCTTGGCAGTGCT[G/T]CTACTCCTCAGGTGCAGCAT	CAGTGTCCTTGGCAGTGCT	TGCTGCACCTGAGGAGTAG	2749	22
WIAF-472	T/C	C	T	GTGTAGGATAGAGTTTTAAA[T/C]ATTGGTATGTGGTGCTAGAG	GTTTTTATGAAGTGTAGGATAGAGTTTTAA	ACTAACTCTAGCACCACATACCAA	19610	8
WIAF-475	C/T	C	T	TGCAACTTCTCAGTTTTATT[C/T]TAACTTTAATGATCTCTGTG	GATATTTTGCAACTTCTCAGTTTTATT	TCTTAAAGCTAGTATAAAAGTCACAGAGA	2753	3
WIAF-478	A/G	A	G	CAGAAAGCCACGGAAGCCAC[A/G]GCCACTAGCCCTGAACCTTG	AAGCCACGGAAGCCAC	AGGTTCAGGGCTAGTGGC	2150	14
WIAF-482	A/G	A	G	AAAAGATGCTCTTAGACTGA[A/G]ATTCATAAAGAGTGCCTCAG	GTCAAAAAGATGCTCTTAGACTGA	GACCTGAGGCACTCTTTATGAA	2152	4
WIAF-483	C/A	A	C	TAGAAAGCACAGGGATGACA[C/A]AAATCACTAAGGAATTCCAC	CTAGAAAGCACAGGGATGACA	GTCTTAGTGGAATTCCTTAGTGATTT	2153	4
WIAF-484	T/C	C	T	CCTTTCCTTCCATTGTCCTC[T/C]CTTGAGATGGGTTGCAAATG	CCTTTCCTTCCATTGTCCTC	CATTTGCAACCCATCTCAAG	2756	4
WIAF-487	C/T	C	T	GGAGACCCCACCTCTCACCA[C/T]TTAGAAAAGGGCATTTCAAG	GACCCCACCTCTCACCA	GCTTGAAATGCCCTTTTCTAA	2155	2

续表

WIAF 命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi 编号	NwLib chrom
WIAF-489	G/A	A	G	TGACTGTGGGACCAGACAAG[G/A]GATGTAGATTGTCACATTCA	TGACTGTGGGACCAGACAAG	GGATTGAATGTGACAATCTACATC	2156	8
WIAF-490	G/C	C	G	TTTGGTGTTTCTCCATCCTA[G/C]GATTCTGCCTTATAATCTTT	TCTTTGGTGTTTCTCCATCCT	GGACAAAGATTATAAGGCAGAATC	2157	8
WIAF-496	A/T	A	T	ACACAGAATCATCAAAGCAC[A/T]ATCTGTGTTTGAGATAAATG	ACACACAGAATCATCAAAGCAC	CAGACTATCATTTATCTCAAACACAGA	19479	8
WIAF-505	G/C	C	G	AATCTCTGCCTTGTATCATG[G/C]TTATCACTGGACACAGCCAC	TCAATCTCTGCCTTGTATCATG	GTGGCTGTGTCCAGTGATAA	2765	13
WIAF-506	T/C	C	T	GGGCTGGACACTGCAGTGAT[T/C]AGGGGCAGGTGTGGGGCAGG	GGCTGGACACTGCAGTGAT	CCCACACCTGCCCCT	2766	16
WIAF-507	A/G	A	G	GAATCTGTGGTTTTCGCCTG[A/G]TAGACCACAGGGCCAATCAC	ATCTGTGGTTTTCGCCTG	TGATTGGCCCTGTGGTCTA	2767	17
WIAF-510	T/C	C	T	TTTATCACAGCAGGGGACAG[T/C]AAGGTTGGCTTCTCTAATGC	TTATCACAGCAGGGGACAG	GGCATTAGAGAAGCCAACCTT	2770	16
WIAF-511	A/C	A	C	CTAATAAGCAAAGACCTGCA[A/C]CCCTGGCTTCCTGACTCCAA	CCTAATAAGCAAAGACCTGCA	GGAGTCAGGAAGCCAGGG	19611	17
WIAF-512	G/A	A	G	TTTGGGTGTGAGCGGATTAT[G/A]TCTGACGCCATGGGTGTTCA	TGGGTGTGAGCGGATTAT	CACCCATGGCGTCAGA	2161	5
WIAF-514	T/G	G	T	CTGTCTCATCCAATAAGCAC[T/G]CATGACCTCAGCCCCATACT	TGTCTGTCTCATCCAATAAGCA	TGGGGCTGAGGTCATG	2771	7
WIAF-523	C/T	C	T	TTGTTGGCACCAGAAAAGCT[C/T]ATGTTCTATGTTATGTCACT	TGTTGGCACCAGAAAAGCT	CAGTATGTACAGTGACATAACATAGAACA	2779	3
WIAF-531	T/G	G	T	TCTTGAACTTTAATAAATAC[T/G]CTTTTTACAAAACACGTGAA	GGTTATGTGTTCTTGAACTTTAATAAATAC	GTACATTCACGTGTTTTGTAAAAAG	2782	1
WIAF-533	T/C	C	T	GCTTCAGCTTAACTGACAGA[T/C]GTTAAAGCTTTCTGGTTAGA	CAGCTTCAGCTTAACTGACAGA	AAACAATCTAACCAGAAAGCTTTAA	2783	12
WIAF-534	G/A	A	G	ACCCAAAACACAGAGACCCC[G/A]TGAGTCTTAGTCAATCCTAT	CCCAAAACACAGAGACCCC	GCCACTATAGGATTGACTAAGACTCA	2169	15
WIAF-537	G/T	G	T	GGCACAGAAGCCAGTCATAC[G/T]TGCTTTAAAATTGACCCAAC	GGCACAGAAGCCAGTCATAC	GGTTGGGTCAATTTTAAAGCA	2170	1
WIAF-538	C/T	C	T	TATTGTGATTGTTTTACATG[C/T]GAAATCTGGCTTCAGAAAGG	CCTTTAGATATATTGTGATTGTTTTACATG	ACCTTTCTGAAGCCAGATTTC	2171	6
WIAF-545	A/C	A	C	TGGACACCAACCAACAGAAT[A/C]CTCCCGTCCTTTGAAATTTC	TGGACACCAACCAACAGAAT	GAAATTTCAAAGGACGGGAG	2790	4
WIAF-548	A/G	A	G	TTCTTCCAAAAACAAAGAAT[A/G]AACATTGGAATAGTCACTTA	GCATTCTTCCAAAAACAAAGAA	CCTTGTAAGTGACTATTCCAATGTT	2792	5
WIAF-550	G/C	C	G	GGTCAGGTCCCCCAGATTGA[G/C]GTCTGAGTGTGGGCAAGTGT	CAGGTCCCCCAGATTGA	CACTTGCCCACACTCAGAC	2793	7
WIAF-552	A/G	A	G	CCTCCAAACATAATTGATTC[A/G]TATCTGCGAGACTTACACTC	CGTACCTCCAAACATAATTGATTC	GCTTGAGTGTAAGTCTCGCAGA	2794	12
WIAF-554	G/A	A	G	AGGGCCTTGCTGGCTACTGG[G/A]TGTTAGTTTGCAGTCCTGTG	GCCTTGCTGGCTACTGG	CACACAGGACTGCAAACTAACA	2796	14
WIAF-559	C/T	C	T	CCTGCATGGTCTTTTCTCTG[C/T]TTTACATCATTGTCATAAAT	CCTGCATGGTCTTTTCTCTG	AGTTGAGATTTATGACAATGATGTAAA	2177	21
WIAF-563	A/C	A	C	GTAGAAAAATTACTATCCAA[A/C]CTGAATTCAGAATAAATAGA	GAACATATTTGTAGAAAAATTACTATCCAA	TCACCTTTCTATTTATTCTGAATTCA	2181	
WIAF-565	C/T	C	T	CAAATAAAGATGGTAGTGAG[C/T]GAACAGAAGAGGTTTCATTG	CAATTTCAAATAAAGATGGTAGTGAG	CAGTTTAGGAGTCAATGAAACCTC	2183	8
WIAF-570	G/A	A	G	AAGAGCGTCCTATGAATCAT[G/A]CATTTGTTCCTGTTATTGCT	CCAAGAGCGTCCTATGAATCA	AACAGCAATAACAGGAACAAATG	2187	7
WIAF-571	C/A	A	C	ATCTAACAGCTGCAGAATGG[C/A]CTTCTTCCTTCCCAGCTTTT	GATCTAACAGCTGCAGAATGG	AAAAGCTGGGAAGGAAGAAG	2802	19
WIAF-575	C/T	C	T	TATTCAGGGAAAACTTTAAT[C/T]TTCTTTGTCTTCTCCAAAAA	GCATTTATTCAGGGAAAACTTTAA	CTGTTTTTGGAGAAGACAAAGAA	2805	5
WIAF-577	T/C	C	T	AAAATGAGTTGAAATAAATG[T/C]AAGTTGAATAATGACACTGT	TCCAAAAGTGATTAGGTGAAAAA	TTTCAACAGTGTCATTATTCAACTT	2189	6
WIAF-579	C/G	C	G	AAATTGCTTTAAGTACTTTA[C/G]GAAGACCTTGACTGTTGGAT	CAAACTTCAAATTGCTTTAAGTACTTTA	AAATCCAACAGTCAAGGTCTTC	19612	3
WIAF-583	G/A	A	G	CCCAGAGAAGCTCTGTCTGC[G/A]CTGCAAAGCCATGGCTGCAG	CCCAGAGAAGCTCTGTCTGC	CAGCCATGGCTTTGCAG	2807	8
WIAF-584	T/C	C	T	AATAAACAATGCAACGTTCC[T/C]GATTTCTAATCTTGGTTCTG	CAAATAAACAATGCAACGTTCC	GCTCTCAGAACCAAGATTAGAAATC	2808	6
WIAF-588	A/G	A	G	TTTGGCTCTGCTACTTGCCA[A/G]ATGAGATTTATTTATGTGGG	TGGCTCTGCTACTTGCCA	GAAACTCCCACATAAATAAATCTCA	19193	11
WIAF-592	A/G	A	G	TTTAACAGACTCAAAAATAT[A/G]GCGAAAACTATCTTTACAAA	GGACCTTAATATTTAACAGACTCAAAAA	AGATAATTTTGTAAAGATAGTTTTCGC	2193	6
WIAF-594	T/G	G	T	CAAAGAGATGGGGTATATAA[T/G]AAAGAACCATGTAAAGATTT	TTGTATCAAAGAGATGGGGTATATAA	AATTAAAGAAATCTTTACATGGTTCTTT	2194	2
WIAF-599	A/C	A	C	TCAAATTTGGACTATGAACA[A/C]GACATAGTTGCTAAGGATAT	GGAATCAAATTTGGACTATGAACA	GTGGAATATCCTTAGCAACTATGTC	19613	21
WIAF-600	A/G	A	G	CAAGGGAGGAACATTTACAG[A/G]GTCCATCTCTGATGTCACCA	GCAAGGGAGGAACATTTACAG	CTGGTGACATCAGAGATGGAC	2197	13
WIAF-601	C/T	C	T	TGTCCCCTGGGAGACTATGG[C/T]AGTGAACACTAAAATCCTAA	TCCCCTGGGAGACTATGG	GCGATTAGGATTTTAGTGTTCAC	2813	5
WIAF-603	C/T	C	T	ATTCAGAAAATTATAGTTCC[C/T]CAAGTTCATGCATAACAGGA	AGTTGAATTATTCAGAAAATTATAGTTCC	TTTCCTGTTATGCATGAACTTG	2198	8
WIAF-606	C/A	A	C	TGTGTCTTCATTGGTTCACT[C/A]TTAAAGTTCTGTTATGCATT	GCTGTGTCTTCATTGGTTCACT	AAGAACAATGCATAACAGAACTTTAA	19614	16
WIAF-608	A/T	A	T	GTTTTCTGTATGTACAACTC[A/T]CCAACCATTAGGATTGTCAC	GGTTGTGTTTTCTGTATGTACAACTC	GAGTGACAATCCTAATGGTTGG	19615	19
WIAF-613	G/T	G	T	GGTGATAATGGGGCCCTGTT[G/T]CTCTGGCATACATATAAAAG	TGATAATGGGGCCCTGTT	CCTTCCTTTTATATGTATGCCAGA	2200	6
WIAF-615	T/A	A	T	ATGAATTCATCTTTCAGTTT[T/A]TAGATCGGATCATGAATTAG	TCTTCATGAATTCATCTTTCAGTTT	GGACTAATTCATGATCCGATCT	2202	8
WIAF-618	A/G	A	G	TAACACCCAGGCCCTTTCCC[A/G]TTATATCCAGGTATGCTACA	CCCAGGCCCTTTCCC	AAGAACTTGTAGCATACCTGGATATAA	2818	15
WIAF-619	C/A	A	C	AACCATGGGTCAACTATGAT[C/A]CCAAAACAGCAGTGTTGTCT	GAGTAACCATGGGTCAACTATGA	TTAGACAACACTGCTGTTTTGG	2203	3
WIAF-620	C/G	C	G	TGGATAAATCATGTGCCCCA[C/G]AGAGCCCCAAAGCTTGATGA	GGATAAATCATGTGCCCCA	ATCAAGCTTTGGGGCTCT	2819	18
WIAF-623	G/C	C	G	TTTGTGCTGCTACGTTGTTA[G/C]AGCAACCCCAGAAAATTAAA	AGTTTGTGCTGCTACGTTGTT	TTTTAATTTTCTGGGGTTGCT	2205	11
WIAF-624	G/A	A	G	ATTAAGAAAAGATAACAAGA[G/A]ATGAATAAATGAGGTAGTGG	GCAAGATATAAAGATTAAGAAAAGATAACA	CAATTCCACTACCTCATTTATTCA	2206	21
WIAF-626	A/G	A	G	TGTCATTTATGCTGCAGTCG[A/G]AATACTTGGAGCCTGGAAGT	TGTCATTTATGCTGCAGTCG	TTACTTCCAGGCTCCAAGTATT	2208	3
WIAF-627	C/T	C	T	TGACAAGCAGCTCATTCAAG[C/T]GGCCCACCATGGCCCTAGGG	TGACAAGCAGCTCATTCAAG	AGGGCCATGGTGGGC	2821	16
WIAF-636	T/C	C	T	CATTTCTGGTAGCAGGTATA[T/C]GGACTCATTTCTTCTTTGCA	TTCCATTTCTGGTAGCAGGT	GATGCAAAGAAGAAATGAGTCC	2214	7
WIAF-638	C/T	C	T	CCCAGAGAAAAATCCAAGAG[C/T]CTTAAACCATATTTTGTGTT	TCCCAGAGAAAAATCCAAGAG	AGTTTCTAAACACAAAATATGGTTTAAG	2216	1
WIAF-640	C/T	C	T	CCTGAGCCCTCAAGAACTCA[C/T]GCCAGCTCAGCCCTACACCA	CCTGAGCCCTCAAGAACTCA	TGTAGGGCTGAGCTGGC	2825	1
WIAF-641	G/A	A	G	ATATTGTCTGCCTTAAAGCA[G/A]TACCCCCCTACCACACACAC	AGAATATTGTCTGCCTTAAAGCA	GGTGTGTGTGGTAGGGGG	2826	12
WIAF-645	G/A	A	G	ACCAGAAACTCTGCGGTGTG[G/A]AGAAAGGACAGTTATGCTGC	ACCAGAAACTCTGCGGTG	CAGCAGCATAACTGTCCTTTCT	2219	11
WIAF-646	C/G	C	G	GGAGAAGGCACCCAGCCATC[C/G]TGACCCAGCGAGGAGCCAAC	AAGGCACCCAGCCATC	GCTCCTCGCTGGGTCA	2829	14
WIAF-648	G/A	A	G	TCCTAGTCTTCTCTTATAGG[G/A]ACCCTGTGATTACACTTAGG	TGACCTTCCTAGTCTTCTCTTATAGG	GCCCTAAGTGTAATCACAGGG	2221	13

续表

WIAF 命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi 编号	NwLib chrom
WIAF-650	T/C	C	T	GAGAGGTCAGATCATATCCA[T/C]AGAAAAACAGCTCTCTTTTA	CGAGAGGTCAGATCATATCCA	GCAAGTAAAAGAGAGCTGTTTTTC	2830	3
WIAF-651	T/C	C	T	GCACAGCACATAGTGGAAAG[T/C]GCTAAGTGTCCTACGAGAGG	AGCACAGCACATAGTGGAAAG	GACCTCTCGTAGGACACTTAGC	2831	3
WIAF-653	C/T	C	T	GGCACCAGCCTTTTTAGAGT[C/T]CCTGGGCAATTTGTGCACTA	GGCACCAGCCTTTTTAGAGT	TGCACAAATTGCCCAGG	2224	1
WIAF-654	A/G	A	G	AGGGACTTAACCTTTGGCCT[A/G]CCTGCCTGGCTGTTTGGCTC	GGGACTTAACCTTTGGCCT	CAAACAGCCAGGCAGG	2832	19
WIAF-657	T/A	A	T	CTCCCACCCCTATATTTTAA[T/A]GCAACTGACAGTTTTGAAGG	CCTCCCACCCCTATATTTTAA	TCCTTCAAAACTGTCAGTTGC	2834	9
WIAF-664	C/T	C	T	GGTTTTCCATCTGTCTTCCA[C/T]AGAGATCTAGGGTGTCTTTG	GGTTTTCCATCTGTCTTCCA	TCCAAAGACACCCTAGATCTCT	2231	4
WIAF-665	A/G	A	G	TTCTCACCATGGGAATCTTG[A/G]TGCAAGTTAGATCCCACCCT	TTCTCACCATGGGAATCTTG	GGGTGGGATCTAACTTGCA	2232	1
WIAF-666	C/T	C	T	AGAAGCCATTTATTTTGCAG[C/T]CTTCAGTCCAAAAAAAGTCA	TGAGAAGCCATTTATTTTGCA	GTTTATTGTTATAAAAATGACCTACAACTT	2835	12
WIAF-671	C/T	C	T	ACAGACTCTGGAGCCACAGC[C/T]GGCTAATACACTGCAATATT	CAGACTCTGGAGCCACAGC	ACATAAAATATTGCAGTGTATTAGCC	2837	7
WIAF-672	T/C	C	T	ACCACCATTATTAAGGAGAG[T/C]ACTAGGAAAAACTACCAAAC	AAACACCACCATTATTAAGGAGAG	GCTGTGTTTGGTAGTTTTTCCT	2838	7
WIAF-673	T/C	C	T	TTTCTACCTCTATTCTCTTA[T/C]TAAACTTTTGGATACATTCC	CTGTTGATTTTCTACCTCTATTCTCTT	GCTTTGGAATGTATCCAAAAGTTT	2236	16
WIAF-678	A/G	A	G	CTTTTGTTATTATGCTCTTA[A/G]TGATTTACAGACTGATGCCA	AGAGTTCACTTTTGTTATTATGCTCTT	TCTGGCATCAGTCTGTAAATCA	2840	7
WIAF-68	C/A	A	C	AGTGAGCCATGAACTGGCTG[C/A]GTTTTCAACCTTTCCTTGGG	TGAGCCATGAACTGGCTG	CCCAAGGAAAGGTTGAAAAC	2473	15
WIAF-685	T/C	C	T	AGCAGAGTACCTTTCTAACT[T/C]ATAAGATTGTGTAGAGGTTG	ATCACAGCAGAGTACCTTTCTAACT	CCTTCCAACCTCTACACAATCTT	2244	3
WIAF-692	G/A	A	G	GTGGAGTGGGCAGATAAAGA[G/A]CCAAGCCCTAGTTTGAGTGA	GTGGAGTGGGCAGATAAAGA	TCACTCAAACTAGGGCTTGG	2247	10
WIAF-693	G/A	A	G	CCCTTCTTAAAGAGATAGTC[G/A]CCAGAGGCAATTCGACTTTC	TTTTCTCCCTTCTTAAAGAGATAGTC	AAAGTCGAATTGCCTCTGG	2248	18
WIAF-694	C/T	C	T	TCTGATGCCCCTGTCCAAGG[C/T]TGTGTCTACACATGAATTTA	TGCCCCTGTCCAAGG	CAATCTCTAAATTCATGTGTAGACACA	2249	14
WIAF-696	A/G	A	G	CCCTGTATCTGTTCAGGCCC[A/G]GAATCGTCACGGCTCACAAC	CCTGTATCTGTTCAGGCCC	TGTGAGCCGTGACGATTC	2250	1
WIAF-700	A/G	A	G	TATTTTAAAATTATCCTTCC[A/G]TGAAATTGGTGAAAGGTCAA	TCACTGTTATTTTAAAATTATCCTTCC	TTTGACCTTTCACCAATTTCA	2252	18
WIAF-706	T/C	C	T	TCTTTATTCTCCTTCTTTCA[T/C]TAATTTTCTTTCACGTTATT	GGCTCTTTATTCTCCTTCTTTCA	CGTTCAGGGTGAGGGAATAA	2256	3
WIAF-712	G/C	C	G	TAACCTGGATATAAACATCT[G/C]ATGGAAGGCTGCACTGGATG	GGTTTCTAACCTGGATATAAACATCT	CCAGTGCAGCCTTCCAT	2259	11
WIAF-713	T/G	G	T	CCATGCCCCTGATAGTTCTG[T/G]GAGCCACCTAAACTCGTTTC	CCATGCCCCTGATAGTTCTG	GGAAACGAGTTTAGGTGGCTC	2260	11
WIAF-715	T/C	C	T	TTTACGTTCCCTGTCTCATC[T/C]TTCTAGGTAATTTGAGTTTC	TGTTTACGTTCCCTGTCTCATC	GGTTGGAAACTCAAATTACCTAGAA	2261	3
WIAF-716	C/T	C	T	AAGAAATGGCTGATACCAAG[C/T]CTGCAGTGAAAAATGCACAT	TGAAGAAATGGCTGATACCAA	ATGTGCATTTTTCACTGCAG	2262	8
WIAF-718	T/C	C	T	CACCACAGTGCACCAAGGAC[T/C]GGACCTGCACTCTATCTTTA	CCACAGTGCACCAAGGAC	GGGTAAAGATAGAGTGCAGGTCC	2264	2
WIAF-721	A/C	A	C	CCAGAGTCTTCTAATAGCAA[A/C]AGCTACTGGAAGCGGCAAGA	GTTACCCAGAGTCTTCTAATAGCAA	TGCCGCTTCCAGTAGCT	19616	3
WIAF-723	G/C	C	G	ATCTGAAGTTTTTGTCTCCA[G/C]AGAAGTCATTTTGTAGGTGT	CAGTATCTGAAGTTTTTGTCTCCA	AGGAACACCTACAAAATGACTTCT	2266	10
WIAF-724	A/G	A	G	GGGAAGGATCCTGTAAAAGG[A/G]TAAATATTGTTTTCCATAAT	GGGAAGGATCCTGTAAAAGG	CATCTTCAATATTATGGAAAACAATATTT	2267	8
WIAF-725	G/T	G	T	TATGAGTCTTGTGAAACAGG[G/T]GTGGGAAGGATCCTGTAAAA	TCAATATGAGTCTTGTGAAACAGG	CTTTTACAGGATCCTTCCCAC	2268	8
WIAF-726	C/T	C	T	ATTAGATTTCTATTTTGATA[C/T]TGATGTTTCTTTCAAGAGGA	TGTGCTTTAAAGTGTGTAAGTATTAATTAG	ATTTCCTCTTGAAAGAAACATCA	2269	6
WIAF-729	C/T	C	T	AGAAGATAACTAGAAAATGC[C/T]GAACAGAAAAATAACCAGAA	GATGCAGAAGATAACTAGAAAATGC	GAACTCTTCTGGTTATTTTTCTGTTC	2271	14
WIAF-730	A/C	A	C	CACTGAAAGACACCATTTAT[A/C]TACCCAAGGGCAGAAAGTAG	CACTTCACTGAAAGACACCATTT	TCTACTTTCTGCCCTTGGGT	2851	20
WIAF-731	C/T	C	T	CATCTTCATGGGCAGGAATT[C/T]CATTTCTGTGTTTCTTAGGG	CATCTTCATGGGCAGGAATT	CAAACCCTAAGAAACACAGAAATG	2272	9
WIAF-736	C/G	C	G	AAGACTATTTGGATAAGTTT[C/G]ACAAAGATGAGAACAGGTCC	AAAAGCAAGACTATTTGGATAAGTTT	TCTAGGACCTGTTCTCATCTTTG	2274	6
WIAF-737	G/A	A	G	AACAGGTCCTAGAACCTCAG[G/A]ATCGAAAGGAAGTTCATCTA	GAGAACAGGTCCTAGAACCTCAG	GGACTAGATGAACTTCCTTTCGA	2275	6
WIAF-741	C/T	C	T	GCTTGGAGATTCTGATTCAG[C/T]GTGCTCAGGCGGGGCTGGAC	CAGCTTGGAGATTCTGATTCAG	GCCCCGCCTGAGCAC	2854	3
WIAF-742	C/T	C	T	CAGTGGAATCATTTACACTA[C/T]CGAAATCAGCAAATGCTAAA	GCCACAGTGGAATCATTTACA	TTTTAGCATTTGCTGATTTCG	19177	2
WIAF-744	T/C	C	T	TCCCAGAGACCACTTCATTC[T/C]TTTTTGGATTATGAAATAGA	TCCCAGAGACCACTTCATTC	CTACTCTTTCTATTTCATAATCCAAAAA	2277	10
WIAF-745	C/T	C	T	AAGTAGCCCAGCTACAGCCT[C/T]GGTGCATCTTAACCCCTCTC	AAGTAGCCCAGCTACAGCCT	GAGAGGGGTTAAGATGCACC	2855	3
WIAF-747	G/A	A	G	AGACACCAAAAATTCCCAAT[G/A]CTCTAAATAGATGGACTCAA	AATAGACACCAAAAATTCCCAAT	GGGGTTGAGTCCATCTATTTAGA	2279	17
WIAF-748	G/A	A	G	AATCGACGAACATCACAGTG[G/A]GCTGTGGTGCCAAGGACGCA	AATCGACGAACATCACAGTG	TCCTTGGCACCACAGC	2280	17
WIAF-750	A/G	A	G	TTTCTAAAATATAATTTGCT[A/G]TAGAGTTCACAGATGAGCAC	GCCTCACAAAGTATTTTCTAAAATATAA	AAAAGTGCTCATCTGTGAACTCT	2857	11
WIAF-751	T/C	C	T	CCAAACAGAATTACTTGGCA[T/C]AGGGTTTCTTAAAACTATTT	ACCAAACAGAATTACTTGGCA	TGCAGAAATAGTTTTAAGAAACCC	2281	2
WIAF-752	C/T	C	T	ACGAATTTGCTGGGGAATCT[C/T]GTTTTTCTTCTTAAGACTTT	CGAATTTGCTGGGGAATCT	TCCCAAAAGTCTTAAGAAGAAAAA	2858	17
WIAF-761	T/C	C	T	CTAGCGTTTGACTTGTGCGG[T/C]GTACTCAAATGGGGGGCAGT	GCGTTTGACTTGTGCGG	TCCCACACTGCCCCC	2864	8
WIAF-762	C/G	C	G	ATGAGAGGATTCCCTCTCTC[C/G]TCCAAGGGAAAGAAGCTTTT	CATGAGAGGATTCCCTCTCTC	AAAAGCTTCTTTCCCTTGGA	2865	6
WIAF-768	C/T	C	T	CAAAGGAAAAGAACAAATTT[C/T]CAAAGACTTGGGGGAGTGAA	GACTTCAAAGGAAAAGAACAAATTT	TCACTCCCCCAAGTCTTTG	2870	1
WIAF-774	A/G	A	G	GGGAGATTTTAGACTGAATC[A/G]TTCTAGAGTATTTGACGACT	GGAGGGAGATTTTAGACTGAATC	AGCTGTAGTCGTCAAATACTCTAGAA	2876	1
WIAF-782	C/T	C	T	CACACAAACTTTTTTAACTC[C/T]GTCAAAAACAACAAGAACAG	TTATATTGTATTGCACACAAACTTTTT	TCATCTGTTCTTGTTGTTTTTGAC	2884	12
WIAF-785	T/C	C	T	TTTGCAATAAACACCATCAT[T/C]CCTGAGTCCACAGATAAGGT	CATTTGCAATAAACACCATCA	GGACCTTATCTGTGGACTCAGG	2887	20
WIAF-786	C/T	C	T	AACACAAACATATCAAGGAT[C/T]GGGCTGGAATCTTTTCCATT	ATGTTTAACACAAACATATCAAGGAT	TGGAAAAGATTCCAGCCC	2888	19
WIAF-791	C/T	C	T	ATTCAAGGCACAAAGCTAAG[C/T]ACATGCAACAATATAGATGA	CAATTCAAGGCACAAAGCTAA	CCAAATCATCTATATTGTTGCATG	2893	15

续表

WIAF 命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi 编号	NwLib chrom
WIAF-796	T/G	G	T	ATCAAACAGTACATGATTAC[T/G]CGGTTTCCAGAAATCTGGAT	CCCAATCAAACAGTACATGATTAC	TCCAGATTTCTGGAAACCG	2898	3
WIAF-798	T/C	C	T	GCAGAAAAAGGCATATTCAA[T/C]TGTCCCATACTAATTTTTGA	CGCAGAAAAAGGCATATTCA	GGTTATTCAAAAATTAGTATGGGACA	2900	14
WIAF-801	G/C	C	G	AGTTTGAATGTATTCCTGAT[G/C]TTTTCCTTTGCCAACTATTC	TGTCAGTTTGAATGTATTCCTGA	TGAATAGTTGGCAAAGGAAAA	2903	18
WIAF-802	A/G	A	G	CACTTGAACTCAAGTCATCA[A/G]TTTTAGGCACAAAGGTTTTA	CCCACTTGAACTCAAGTCATCA	AAACTAAAACCTTTGTGCCTAAAA	2904	15
WIAF-806	G/C	C	G	TGGCTTTAGAATCAAATGGG[G/C]TGACTTTTTCCCCTGTTGGT	TGGCTTTAGAATCAAATGGG	CCAACAGGGGAAAAAGTCA	2908	7
WIAF-807	T/C	C	T	TGAAATGGGCACTTAGCGAA[T/C]ACTTGTGGACCACAAGACTT	GAAATGGGCACTTAGCGAA	CAGACAAGTCTTGTGGTCCAC	2909	7
WIAF-814	C/T	C	T	TGGTGATGCAATTGAAATAA[C/T]ATTGTAAGTTAATGTAATTG	TCTGGTGATGCAATTGAAATAA	GCTGCCAATTACATTAACTTACAA	2915	14
WIAF-815	C/T	C	T	AGGAGGAGTCCCTCATGGAT[C/T]GCGGTATTGGTTGGTTGTGG	GGAGGAGTCCCTCATGGAT	CACAACCAACCAATACCGC	2916	1
WIAF-816	T/A	A	T	TACCCACCACACTACCCTGT[T/A]AAAATCTTAACATTGTGATG	ACCCACCACACTACCCTGT	GAGGCATCACAATGTTAAGATTTT	2917	3
WIAF-818	G/A	A	G	ACATTCCCATCTGTCTTGCA[G/A]GGAGGGATCTTGGTCGGCTT	CATTCCCATCTGTCTTGCA	CCGACCAAGATCCCTCC	2919	12
WIAF-819	T/C	C	T	GGGATAGAGGACAGAGTGTT[T/C]GTTGATTTTTCGTTTCGGTT	GGAGGGATAGAGGACAGAGTGTT	AAACCGAAACGAAAAATCAA	2920	8
WIAF-820	C/T	C	T	TATAACCAAGAAGCCTTCAG[C/T]AGAGCAAGTCTGAGCCAGAG	ATCTTATAACCAAGAAGCCTTCAG	CTCTGGCTCAGACTTGCTCT	2921	11
WIAF-821	T/C	C	T	CATTCCAAGGCTCTCTAACA[T/C]GAGTGTCTGCAGCCCCATTC	GACATTCCAAGGCTCTCTAACA	TGGGGCTGCAGACACTC	2922	17
WIAF-823	C/A	A	C	TAGTGATAGTAGAGGACTCA[C/A]CCTGCACGTGCACCTTTCAT	CCAATTTCTAGTGATAGTAGAGGACTCA	AAGGTGCACGTGCAGG	2924	14
WIAF-827	G/A	A	G	TAATGAAACATCTGCGAAAA[G/A]AAGTGTGGGAATCACCTCAT	AGATAATGAAACATCTGCGAAAA	GATGAGGTGATTCCCACACTT	2928	7
WIAF-828	G/A	A	G	CAACTAGCTTGCTTTTTGTC[G/A]TTTGGAAGACTACCATTATT	AATCAACTAGCTTGCTTTTTGTC	TTTGAATAATGGTAGTCTTCCAAA	2929	7
WIAF-829	A/T	A	T	ATATAAAATTACTAAATTAA[A/T]GTCTTAAAAGAAAATATAAC	CCTTTATTTTCCCAAATATAAAATTACT	GTCACCATGTTATATTTTCTTTTAAGAC	2930	13
WIAF-832	C/A	A	C	TGCACAACCAGGGCAAAATA[C/A]TGCTGGATTAACCCATGCTA	GCACAACCAGGGCAAAATA	GCATGGGTTAATCCAGCA	2933	8
WIAF-834	A/C	A	C	AACCCTTTCATCCATTCAGC[A/C]AATTTAAAACTCTTGCCAAG	ACCCTTTCATCCATTCAGC	TGATACTTGGCAAGAGTTTTAAATT	2935	1
WIAF-835	G/A	A	G	TTATTCCAAGAATGGGAAGC[G/A]CATTTTCATTGGCTTGAATG	TTTATTCCAAGAATGGGAAGC	TCATTCAAGCCAATGAAAATG	2936	14
WIAF-838	T/C	C	T	CATGGGCCCTTGGCACTATG[T/C]CTACTCTGCCTGACGGATAA	GGGCCCTTGGCACTATG	ACTTATCCGTCAGGCAGAGTAG	2939	3
WIAF-839	T/C	C	T	TTGACACAAAGTTCTAACTT[T/C]TTGTTAAAAATCTCTGGCTT	GGGCTTGACACAAAGTTCTAA	GGAAAGCCAGAGATTTTTAACAA	2940	19
WIAF-841	T/A	A	T	GCTCTCTGTCCCTGGAGGTA[T/A]GCAAGAGGGTGGAGAAGTCT	GCTCTCTGTCCCTGGAGGTA	GACTTCTCCACCCTCTTGC	2942	11
WIAF-842	G/A	A	G	GAACAGACTGGAGTGAGAAC[G/A]GGTTCCACCACCAAGCCCCT	TGGAACAGACTGGAGTGAGAA	GGCTTGGTGGTGGAACC	2943	11
WIAF-843	C/T	C	T	TTGATTTGCTTCGTTCAAAG[C/T]GCTTAGAATGGAAGATTTAG	GTTGATTTGCTTCGTTCAAAG	TCAAACTAAATCTTCCATTCTAAGC	2944	22
WIAF-844	A/C	A	C	CTAAGGAAAACTTTCTCCAT[A/C]AAGCTGCCTGCTGTGCACGT	GCCACTAAGGAAAACTTTCTCC	GCACAGCAGGCAGCTT	2945	6
WIAF-845	A/C	A	C	TACAGGAAGAGTTGTCTCAT[A/C]AGGTGCCACTAAGGAAAACT	CATGTACAGGAAGAGTTGTCTCA	AAGTTTTCCTTAGTGGCACCT	2946	6
WIAF-846	C/T	C	T	GGGAGCCCTAGTTGCAGTAA[C/T]GTGTCATAATAAATAATTGC	GGGAGCCCTAGTTGCAGTAA	CCTGAATATGCAATTATTTATTATGACA	2947	1
WIAF-847	C/G	C	G	TCATTCTCATTTAAAATTGT[C/G]AAATAAGTCAGAAAAAGATC	CAAGAATCATTCTCATTTAAAATTGT	GGAGATATTGATCTTTTTCTGACTTATTT	2948	3
WIAF-852	A/G	A	G	CTAAGTTATTTTAATATTCC[A/G]GGATTTAATTTCTTCCTAGT	TGGGGAATGGGATATCCA	GATTGAACTAGGAAGAAATTAAATCC	2953	5
WIAF-856	T/C	C	T	AACAACAAAGCTGAAACTAA[T/C]GAGTGAGCATAGCAAGATTA	AATCAACAACAAAGCTGAAACTAA	GCCTGTAATCTTGCTATGCTCA	2957	12
WIAF-858	T/G	G	T	GATGATCCCCGACAGCAAAA[T/G]GTTTCCTTTCTGAGGCTGCC	GATCCCCGACAGCAAAA	GCAGCCTCAGAAAGGAAAC	2959	1
WIAF-861	T/C	C	T	ATTCAATGGTTTTTACTCTA[T/C]TGTCAAAGCTGGGCAACTAT	TGTAAACAGTAGCATTCAATGGTT	AGTTGCCCAGCTTTGACA	2962	2
WIAF-865	A/C	A	C	ACACCGACATACATATCTCA[A/C]GTAGAATTAGCTATACTGCA	CTCACACCGACATACATATCTCA	CCTACTACAATGAAGTTAGTATGCAGTAT	2965	8
WIAF-866	G/A	A	G	TATGGCCATTTTAATAAGTG[G/A]TATGCTTCTGAACACCTGC	TTCATATGGCCATTTTAATAAGTG	GGCAGGTGTTCAGAAAGCAT	2966	1
WIAF-867	A/G	A	G	TTAGGGGTGTTGACACACCA[A/G]TTTTGAGTGTACTGTGCCTG	AGGGGTGTTGACACACCA	CCAGGCACAGTACACTCAAAA	2967	8
WIAF-874	A/G	A	G	TTTACAGTCATTTATTTACC[A/G]GTCATGAATTCATTAAAAAC	GGGCTTTTACAGTCATTTATTTACC	TGTGGTTTTTAATGAATTCATGAC	2974	14
WIAF-878	A/G	A	G	ACCAACATTGTCCACCGGGC[A/G]TTGAGAATACAATATTGAAG	CATTGTCCACCGGGC	TGACTCTTCTTCAATATTGTATTCTCAA	2978	2
WIAF-879	G/A	A	G	TTATGAAATATTAAACCTCT[G/A]CTTCTCAGGAGTGACACTAA	AATCTGCAATGAAATCAGTTATGA	CATGATTAGTGTCACTCCTGAGAA	2979	12
WIAF-88	A/G	A	G	CCGTCGATTCTGAATATCCC[A/G]TGGCGGCATATGCAAAGGAA	CCGTCGATTCTGAATATCCC	TTGCATATGCCGCCA	2490	15
WIAF-881	G/A	A	G	AAGAACTGTGAACTATGAAC[G/A]TCTCAGCCTAGAAGGATAAT	CCTGAAGAACTGTGAACTATGAAC	CACATTATCCTTCTAGGCTGAGA	2981	6
WIAF-883	T/C	C	T	AAGGTCAGGGCAACATAGGA[T/C]TGTGACAGCACCACTCGGAC	AGGTCAGGGCAACATAGGA	CCGAGTGGTGCTGTCACA	2983	1
WIAF-889	A/G	A	G	TCGCGTGGAACCTGAAACAC[A/G]GACGCCTTTCTTCCAAGAAG	GCGTGGAACCTGAAACAC	TTCTTGGAAGAAAGGCGTC	2989	16
WIAF-893	C/T	C	T	GCGGAAGACAGTGAGCTGTT[C/T]GAGCTGGATTATTGCCTCAA	GCGGAAGACAGTGAGCTGTT	TTGAGGCAATAATCCAGCTC	2993	19
WIAF-894	C/T	C	T	ACTCATCATCCCAGATTATT[C/T]TGAAGTGGAAACCACCCTCC	TGTCTAACTCATCATCCCAGATT	GAGGGTGGTTTCCACTTCA	2994	19
WIAF-895	G/A	A	G	ATCTATATTGTACCAGAACT[G/A]TTTATTTCACCCCATCAAGT	GGACAGAAGATCTATATTGTACCAGAA	CTTATACTTGATGGGGTGAAATAAA	2995	4
WIAF-896	C/A	A	C	AATCATATAGCCAAGGGACT[C/A]GGAATTTTGGCTGCTTCAAG	TCATAATCATATAGCCAAGGGAC	TGAAGCAGCCAAAATTCC	2996	3
WIAF-899	C/T	C	T	AGTGTATTTTTCCTGTTCCA[C/T]GAAGAGGACTTTTTGTTCAC	GGAGTGTATTTTTCCTGTTCCA	AATTGTGAACAAAAAGTCCTCTTC	2999	21
WIAF-901	T/C	C	T	CCTGAGACATGTGGTGGCTG[T/C]GAGGGAGAAGGACCCCCAAC	TGAGACATGTGGTGGCTG	GGGGGTCCTTCTCCCTC	3001	10
WIAF-902	C/G	C	G	TTAAAATGATTCAGCATTTA[C/G]ACTTTAAAAATTACCTCAAT	TGAGGTTAAAATGATTCAGCATTT	AGGAACATTGAGGTAATTTTTAAAGT	3002	7
WIAF-903	T/G	G	T	TAAGTCGGGAGTGCTGATTG[T/G]TCGGGTCCAAGATAAAATCT	AGTCGGGAGTGCTGATTG	AAGATTTTATCTTGGACCCGA	3003	4
WIAF-906	C/T	C	T	AAGGCCTAGAGGAGAGATTA[C/T]AATGAACGTAAATAATTCAA	ACAAAGGCCTAGAGGAGAGATT	TGCCTTGAATTATTTACGTTCA	3006	3
WIAF-907	A/G	A	G	GTGCAAATGACAAAGCCCAA[A/G]AGAACAGAGGATCAAATAAG	GCAAATGACAAAGCCCAA	TCAATCTTATTTGATCCTCTGTTCT	3007	18
WIAF-908	C/G	C	G	AGGAAGTCTGGGGATTCCTA[C/G]AGGGGACATATCACACATAT	AGGAAGTCTGGGGATTCCTA	TTAGAATATGTGTGATATGTCCCC	3008	17
WIAF-911	G/A	A	G	CAAGCGAGAGGGAGGCAAGC[G/A]GAGGGATGGTTTATCATAGC	CGAGAGGGAGGCAAGC	GGCTATGATAAACCATCCCTC	3011	13
WIAF-914	T/C	C	T	ACTGAGCGATGACACCACAC[T/C]TTGTTTGGACATTTAAATTC	TGAGCGATGACACCACAC	GAGTGAATTTAAATGTCCAAACAA	3014	2

续表

WIAF命名	等位基因	等位基因 A	等位基因 B	芯片中的位置	上游引物	下游引物	ncbi编号	NwLib chrom
WIAF-915	G/A	A	G	CGCTCTCTTTTCTCCCAAGC[G/A]AAACCAAATGCGCCCCTTCA	CGCTCTCTTTTCTCCCAAGC	GGGGCGCATTTGGTTT	3015	6
WIAF-916	T/C	C	T	TATTCTGCTATTCCTGCCAT[T/C]ACCGCATCCTTCATGGTAGA	AATTATTCTGCTATTCCTGCCA	ACCATGAAGGATGCGGT	3016	6
WIAF-917	A/G	A	G	GAGACACTACTCTTCCTTCA[A/G]GACTATTTCATTCTGACTAT	GCTGAGACACTACTCTTCCTTCA	TTCAATGTATTTATTCACTTATAGTCAGAA	3017	3
WIAF-918	T/C	C	T	ATCATTTTAGCGTAAATACC[T/C]GAATAACCCATAGTTACAGA	GCATTTTAACTATCATTTTAGCGTAAAT	CCAATTCTGTAACTATGGGTTATTC	3018	17
WIAF-919	A/G	A	G	CACTCCATCCTGATGCCCCA[A/G]GTTATCCACAGCCTCCTTCC	CCATCCTGATGCCCCA	GGAAGGAGGCTGTGGATAAC	3019	4
WIAF-922	A/C	A	C	GTCATTACACACTGAGCAAC[A/C]AAACAAAGGTGTTGAATCCT	TCAGTCATTACACACTGAGCAAC	AAGAGGATTCAACACCTTTGTTT	3022	6
WIAF-923	C/T	C	T	CAGGGGTGACGTATGTAGAA[C/T]GCTTAGGGTGTCCTCCCCAC	GCAGGGGTGACGTATGTAGAA	GGGGAGGACACCCTAAGC	3023	19
WIAF-924	G/T	G	T	CTACCTGAAACAGGGATGCC[G/T]TTCTCGGTACTATGTTTAAT	CCTGAAACAGGGATGCC	GCACAATTAAACATAGTACCGAGAA	3024	16
WIAF-931	T/C	C	T	CAGAGGAACAGCCAAGAGAT[T/C]TTACCGTGGTCTTACTAAAG	CAGAGGAACAGCCAAGAGAT	GATATGTACTTTAGTAAGACCACGGTAA	3031	4
WIAF-933	G/A	A	G	GTGGTTTTCAATACTAAACA[G/A]TGTAAACAATGCAAATATTT	AGCGTGGTTTTCAATACTAAACA	CTGTATTTATTGTTAAATATTTGCATTGTT	3033	2
WIAF-939	A/T	A	T	CCTGAAGGAATCTGGGAATT[A/T]GCCCTGGCCTGAAAGTGGCC	CCTGAAGGAATCTGGGAATT	ACTTTCAGGCCAGGGC	3038	10
WIAF-944	A/G	A	G	AGGATGAAAACTGATACACC[A/G]GTTACTACTTACTCTTCACT	AACAGGATGAAAACTGATACACC	CAGTTTGAAGAGTGAAGAGTAAGTAGTAA	3043	1
WIAF-951	T/C	C	T	GCTGGGTAGTTCTAGAACTT[T/C]AGAAGCTCCATCTTTTAATG	TTCCAGCTGGGTAGTTCTAGAA	AAAAACATTAAAAGATGGAGCTTC	3050	9
WIAF-952	T/G	G	T	TACATGTTTTATATCACTTA[T/G]TTATCTCAACAATCTTGAAA	CAGAACTAAGCTTTACATGTTTTATATCAC	CCCTTTCAAGATTGTTGAGATAA	3051	1
WIAF-953	G/A	A	G	GCTGCTGTCCCAGTGCCACA[G/A]TGGTCTAGCCTCATGGCAGA	GCTGTCCCAGTGCCACA	TGCCATGAGGCTAGACCA	3052	20
WIAF-955	C/T	C	T	CTGCTTAGGTTTGTTCTTCC[C/T]AGCATATTCAGCTATAATCA	GCTGCTTAGGTTTGTTCTTCC	TGTAGGTGATTATAGCTGAATATGCT	3054	2
WIAF-960	G/A	A	G	GCTCTTGGCTTTCAAAATCA[G/A]TACAGACAGATAAGAGCTTT	GCTCTTGGCTTTCAAAATCA	CGAAATACTTAAAGCTCTTATCTGTCTGT	3059	9
WIAF-961	C/G	C	G	GGACTTTCAGTTGACAGATT[C/G]TCTCCTTACCTAACTCTAGT	AAGGGACTTTCAGTTGACAGATT	TCAGGAGAGTCTGACTAGAGTTAGGT	3060	7
WIAF-963	G/A	A	G	ACAGTACAACTTGCAAGCAC[G/A]GGGAGAAAACCTAGGTGAAA	TGTTACAGTACAACTTGCAAGCAC	TTTTCACCTAGGTTTTCTCCC	3062	17
WIAF-966	A/G	A	G	AGTGTGACCTCCTGTGACCC[A/G]TGAATGTGCCTCCAAGCGGC	CAGTGTGACCTCCTGTGACC	CGCTTGGAGGCACATTCA	3065	10
WIAF-967	T/C	C	T	TGGAAATGGAAAGAGTGACA[T/C]CCTTGTCCTGTTCTGAACTT	GTGGAAATGGAAAGAGTGACA	CAGTAAGTTCAGAACAGGACAAGG	3066	6
WIAF-971	G/A	A	G	CTCAAAGCTCTCAGGACTGG[G/A]GCTAGGGTTTAAGGAAGGCT	CTCAAAGCTCTCAGGACTGG	AAGCCTTCCTTAAACCCTAGC	3070	11
WIAF-973	A/G	A	G	ATGAATAGCAGCTTTTAGGG[A/G]TTATATCATGAGGTACAAAT	TCATGAATAGCAGCTTTTAGGG	GCCTCTTTATTTGTACCTCATGAT	3072	13
WIAF-974	C/G	C	G	TCCACAGGTCACACTTGTCA[C/G]CAGCAAGTATAAACAAAGTG	TCCACAGGTCACACTTGTCA	ACCCACTTTGTTTATACTTGCTG	3073	16
WIAF-975	G/C	C	G	AATTCCACAACCACCTTGAC[G/C]AATGCTTGCCAAGCTGTTTT	AATTCCACAACCACCTTGAC	AAACAGCTTGGCAAGCATT	3074	9
WIAF-977	T/C	C	T	GAAGTGTCATCCTGACTGAC[T/C]GTCCCTGCAGTGCCCATGCG	CGAAGTGTCATCCTGACTGAC	TGGGCACTGCAGGGAC	3076	5
WIAF-978	A/G	A	G	AACTCTAACATCGTCTATAG[A/G]ACCATTTCCCGTCTCCAGTT	CCATAGTTGAACTCTAACATCGTC	TGGAGACGGGAAATGGT	3077	19
WIAF-979	G/A	A	G	TCTGTGCTCTTCATTCTCAC[G/A]GGCCCGCAACCCCTCCTCTC	GCTCTGTGCTCTTCATTCTCA	AGGAGGGGTTGCGGG	3078	3
WIAF-980	G/T	G	T	AAAGACATGACTACCAACGC[G/T]GGCCCCTTGCACCCATACTG	CAAAGACATGACTACCAACGC	TGGGTGCAAGGGGCC	3079	X
WIAF-982	C/T	C	T	TTTAAAAATGATCATTTAAC[C/T]TCTTTGAACTACAGCCTGAA	CTTCCTATTGTTTTAAAAATGATCATTTAA	GGATTCAGGCTGTAGTTCAAAGA	3081	6
WIAF-983	G/C	C	G	GCCATGGAAGGAAAGTATGC[G/C]GTGTTTTAGGGAGAGCTGGC	GCCATGGAAGGAAAGTATGC	GCCAGCTCTCCCTAAAACAC	3082	7
WIAF-984	C/T	C	T	CTTTCCTCATCTGCTGGTGT[C/T]TTCCTCAGAGCTTTAATGTC	CTTTCCTCATCTGCTGGTGT	CGGACATTAAAGCTCTGAGGA	3083	1
WIAF-985	T/A	A	T	TCTAGGCTGTGGGGAACCTC[T/A]GGTGCCTTACAACTCCAACT	GGCTGTGGGGAACCTC	GTAGTTGGAGTTGTAAGGCACC	3084	6
WIAF-988	C/T	C	T	AACAATTCAGTATCATTATG[C/T]TGTAGATTTCAGATGTAGGT	TGATGTTTACATAACAATTCAGTATCAT	GACGACCTACATCTGAAATCTACA	3087	9
WIAF-989	A/G	A	G	AGGTTCTGATGGAAGCATCA[A/G]TGATGGATTTGGCTTTTCCA	GGTTCTGATGGAAGCATCA	GGAAAAGCCAAATCCATCA	3088	8
WIAF-991	A/G	A	G	CCATGGGCCCTGAGCGTCCT[A/G]CCCCGAATTCACGAGGCTGA	GGCCCTGAGCGTCCT	CCTCGTGAATTCGGGG	3090	11
WIAF-992	A/G	A	G	GCAGCTGGTTCCTCCAGGGA[A/G]TTGGCCCCGAAGCTGGCTCA	GCTGGTTCCTCCAGGGA	CAGCTTCGGGGCCAA	3091	11
WIAF-993	T/A	A	T	TGGAAAGTCCCTAGGTGTGT[T/A]TGAAACAATCTGGGTATTAC	TGGAAAGTCCCTAGGTGTGT	CCTGTAATACCCAGATTGTTTCA	3092	8
WIAF-996	C/G	C	G	TTCGACCCTACTTAGACCCT[C/G]GCACACAAAGGTTGATCAAA	TGTTCGACCCTACTTAGACCC	TTTGATCAACCTTTGTGTGC	3095	8
WIAF-999	C/T	C	T	AGCCCTTTGTCCATTGTTGA[C/T]TGTGAATAATTGGCAGCAGG	GCCCTTTGTCCATTGTTGA	CCTGCTGCCAATTATTCACA	3098	2

表 6-7 Affymetrix HuSNP 芯片中的 SNP 在图谱上的位置

SNP名称	染色体位置	遗传距离/cM	SNP名称	染色体位置	遗传距离/cM
WIAF-768	1	0	WIAF-4066	1	195.79
WIAF-984	1	0.32	WIAF-2818	1	211.3
WIAF-3800	1	6.49	WIAF-1121	1	213.43
WIAF-2674	1	9.55	WIAF-3460	1	213.64
WIAF-4507	1	28.87	WIAF-2644	1	213.73
WIAF-2292	1	39.96	WIAF-4543	1	219.17
WIAF-1577	1	59.56	WIAF-4237	1	219.55
WIAF-1677	1	64.45	WIAF-1018	1	219.64
WIAF-1970	1	68.96	WIAF-413	1	220.92
WIAF-944	1	71.08	WIAF-3263	1	222.12
WIAF-4027	1	76.62	WIAF-3215	1	223.18
WIAF-3095	1	80.12	WIAF-2441	1	226.74
WIAF-1504	1	80.13	WIAF-2122	1	227.43
WIAF-3825	1	82.31	WIAF-270	1	240.47
WIAF-4561	1	85.6	WIAF-3247	1	249.22
WIAF-1934	1	85.69	WIAF-2636	1	249.5
WIAF-774	1	88.26	WIAF-2078	1	252.69
WIAF-2162	1	89.36	WIAF-3122	1	254.33
WIAF-2912	1	89.68	WIAF-1470	1	255.33
WIAF-640	1	99.2	WIAF-3762	1	257.23
WIAF-1768	1	101.27	WIAF-3529	1	257.96
WIAF-638	1	103.93	WIAF-952	1	258.12
WIAF-3293	1	110.98	WIAF-2149	1	260.25
WIAF-846	1	113.86	WIAF-221	1	260.43
WIAF-653	1	115.56	WIAF-2847	1	260.59
WIAF-3233	1	122.36	WIAF-883	1	260.96
WIAF-2921	1	123.5	WIAF-234	1	263.39
WIAF-3891	1	136.07	WIAF-696	1	266.61
WIAF-1854	1	137.58	WIAF-1779	1	266.98
WIAF-3042	1	138.28	WIAF-4501	1	267.83
WIAF-1748	1	144.48	WIAF-3430	1	271.03
WIAF-4050	1	147.51	WIAF-1013	1	286.03
WIAF-3232	1	149.02	WIAF-1014	1	286.03
WIAF-4566	1	149.07	WIAF-2014	1	286.24
WIAF-1842	1	149.13	WIAF-815	1	287.55
WIAF-1642	1	149.41	WIAF-531	1	291.18
WIAF-2866	1	151.2	WIAF-4029	1	292.32
WIAF-1543	1	155.14	WIAF-200	1	293.8
WIAF-3565	1	158.36	WIAF-1610	1	293.82
WIAF-1417	1	158.57	WIAF-2072	1	293.83
WIAF-4568	1	158.59	WIAF-1660	1	299.35
WIAF-2007	1	165.68	WIAF-2930	2	6.44
WIAF-3798	1	167.52	WIAF-4535	2	10.08
WIAF-3632	1	169.21	WIAF-2853	2	12.83
WIAF-2676	1	170.43	WIAF-3001	2	15.91
WIAF-3038	1	172.82	WIAF-3168	2	16.31
WIAF-858	1	183.11	WIAF-933	2	37.58
WIAF-455	1	183.24	WIAF-1004	2	46.62
WIAF-2182	1	188.16	WIAF-3348	2	58.24
WIAF-3921	1	190.89	WIAF-2638	2	61.71

SNP名称	染色体位置	遗传距离/cM	SNP名称	染色体位置	遗传距离/cM
WIAF-955	2	65.3	WIAF-594	2	233.18
WIAF-3449	2	67.06	WIAF-1942	2	233.36
WIAF-2514	2	70.74	WIAF-1763	2	241.65
WIAF-2265	2	72.83	WIAF-2670	2	241.89
WIAF-3861	2	74.65	WIAF-2671	2	242
WIAF-3250	2	75.08	WIAF-878	2	242.63
WIAF-2649	2	83.34	WIAF-4538	2	246.79
WIAF-2318	2	88.76	WIAF-2686	2	247.66
WIAF-2167	2	98.79	WIAF-1031	2	250.33
WIAF-3538	2	101.69	WIAF-1023	2	251.2
WIAF-1146	2	105.58	WIAF-3216	2	258.47
WIAF-4565	2	105.63	WIAF-3785	2	261.67
WIAF-2635	2	106.02	WIAF-301	2	264.58
WIAF-487	2	108.66	WIAF-3462	2	265.61
WIAF-2130	2	111.41	WIAF-3602	2	271.47
WIAF-1665	2	113.53	WIAF-3143	2	273.32
WIAF-4500	2	113.83	WIAF-1930	2	273.48
WIAF-2449	2	114.03	WIAF-1929	2	273.95
WIAF-2739	2	116.92	WIAF-1700	2	274.68
WIAF-3894	2	123.09	WIAF-4596	2	282.25
WIAF-718	2	125.38	WIAF-3473	2	284.3
WIAF-4095	2	127.42	WIAF-1762	2	284.65
WIAF-4549	2	127.42	WIAF-685	3	11.82
WIAF-2173	2	130.39	WIAF-1793	3	13.84
WIAF-2105	2	131.25	WIAF-3235	3	19.6
WIAF-1964	2	133.35	WIAF-3008	3	21.91
WIAF-751	2	133.73	WIAF-838	3	23.38
WIAF-4562	2	145.18	WIAF-3865	3	35.2
WIAF-3442	2	145.24	WIAF-906	3	35.22
WIAF-3443	2	145.24	WIAF-4570	3	37.71
WIAF-305	2	157.43	WIAF-2299	3	38.48
WIAF-3663	2	157.53	WIAF-2703	3	41.67
WIAF-3556	2	167.83	WIAF-3380	3	50.08
WIAF-3667	2	171.72	WIAF-741	3	56.82
WIAF-1886	2	171.8	WIAF-1789	3	59.6
WIAF-4571	2	171.95	WIAF-1024	3	59.63
WIAF-4502	2	175.64	WIAF-721	3	61.6
WIAF-3616	2	192.65	WIAF-475	3	63.07
WIAF-2333	2	193.89	WIAF-2841	3	65.28
WIAF-861	2	196.04	WIAF-796	3	65.82
WIAF-3271	2	198.14	WIAF-816	3	67.71
WIAF-337	2	198.25	WIAF-2952	3	70.15
WIAF-2701	2	198.59	WIAF-1938	3	70.68
WIAF-2135	2	209.7	WIAF-2626	3	70.68
WIAF-223	2	210.13	WIAF-2197	3	70.91
WIAF-2376	2	215.78	WIAF-3733	3	71.76
WIAF-2080	2	216.56	WIAF-3534	3	71.91
WIAF-3338	2	217.67	WIAF-1489	3	80.52
WIAF-4541	2	223.88	WIAF-4098	3	100.66
WIAF-2827	2	225.63	WIAF-715	3	103.38
WIAF-2869	2	226.09	WIAF-1702	3	104.3

续表

SNP名称	染色体位置	遗传距离/cM	SNP名称	染色体位置	遗传距离/cM
WIAF-1703	3	104.3	WIAF-979	3	232.49
WIAF-3407	3	105.31	WIAF-1007	3	239.21
WIAF-745	3	105.49	WIAF-3566	3	239.23
WIAF-3709	3	105.88	WIAF-3858	4	0
WIAF-2828	3	108.3	WIAF-2091	4	0.81
WIAF-3151	3	115.31	WIAF-1593	4	2.33
WIAF-2871	3	115.72	WIAF-1958	4	2.33
WIAF-3112	3	116.96	WIAF-3225	4	3.05
WIAF-3103	3	117	WIAF-903	4	16.19
WIAF-2804	3	120.63	WIAF-2196	4	16.24
WIAF-456	3	121.56	WIAF-4142	4	17.08
WIAF-896	3	123.01	WIAF-3372	4	21.08
WIAF-1714	3	129.55	WIAF-316	4	25.1
WIAF-3387	3	132.17	WIAF-3750	4	25.18
WIAF-2572	3	134.38	WIAF-3885	4	34.14
WIAF-1629	3	138.74	WIAF-2788	4	38.37
WIAF-3902	3	144.32	WIAF-3459	4	39.06
WIAF-2081	3	145.29	WIAF-2325	4	48
WIAF-1473	3	145.84	WIAF-3923	4	54.82
WIAF-3176	3	146.66	WIAF-1925	4	57.8
WIAF-3909	3	146.76	WIAF-3412	4	57.85
WIAF-2116	3	146.78	WIAF-4522	4	59.56
WIAF-3664	3	147.05	WIAF-931	4	69.94
WIAF-2922	3	147.96	WIAF-4527	4	80.2
WIAF-2803	3	148.6	WIAF-1032	4	80.23
WIAF-2074	3	149.43	WIAF-482	4	82.65
WIAF-393	3	149.47	WIAF-483	4	82.65
WIAF-394	3	149.47	WIAF-919	4	82.72
WIAF-395	3	149.47	WIAF-2288	4	88.26
WIAF-343	3	149.67	WIAF-3047	4	94.02
WIAF-2730	3	149.7	WIAF-2114	4	97.93
WIAF-2059	3	158.75	WIAF-302	4	98.19
WIAF-1749	3	167.06	WIAF-3361	4	98.2
WIAF-650	3	169.09	WIAF-484	4	101.48
WIAF-651	3	169.09	WIAF-1302	4	107.23
WIAF-291	3	169.24	WIAF-2765	4	109.61
WIAF-1834	3	174.93	WIAF-2920	4	110.22
WIAF-1045	3	182.85	WIAF-2592	4	113.52
WIAF-313	3	183.16	WIAF-2791	4	114.57
WIAF-2753	3	183.28	WIAF-1753	4	117.5
WIAF-2749	3	183.32	WIAF-2964	4	120.29
WIAF-3300	3	187.11	WIAF-895	4	127.08
WIAF-626	3	189.38	WIAF-2305	4	129.28
WIAF-1819	3	201.32	WIAF-3818	4	132.94
WIAF-4251	3	202.57	WIAF-3821	4	133.02
WIAF-1786	3	206.68	WIAF-3086	4	135.86
WIAF-917	3	206.97	WIAF-664	4	145.01
WIAF-1906	3	208.4	WIAF-1475	4	154.06
WIAF-249	3	212.96	WIAF-4222	4	155.26
WIAF-250	3	212.96	WIAF-3265	4	160.82
WIAF-2422	3	226.3	WIAF-2936	4	169.1

SNP名称	染色体位置	遗传距离/cM	SNP名称	染色体位置	遗传距离/cM
WIAF-3144	4	172.39	WIAF-1690	6	2.47
WIAF-3418	4	185.32	WIAF-3766	6	3.69
WIAF-3118	4	187.49	WIAF-4574	6	6.24
WIAF-2907	4	193.08	WIAF-967	6	18.6
WIAF-3256	4	202.96	WIAF-1567	6	18.69
WIAF-2975	4	204.27	WIAF-2546	6	18.7
WIAF-3557	4	211.48	WIAF-881	6	19.02
WIAF-4594	4	221.12	WIAF-2892	6	26.63
WIAF-3262	5	21.03	WIAF-2597	6	39.36
WIAF-977	5	29.05	WIAF-2977	6	39.74
WIAF-852	5	29.06	WIAF-613	6	40.26
WIAF-4189	5	47.05	WIAF-3166	6	44.42
WIAF-3342	5	55.89	WIAF-3485	6	48.59
WIAF-1046	5	61.4	WIAF-3805	6	48.64
WIAF-3575	5	65.38	WIAF-4257	6	49.29
WIAF-2999	5	66.08	WIAF-1084	6	49.54
WIAF-601	5	66.4	WIAF-258	6	49.79
WIAF-2705	5	67.14	WIAF-1009	6	49.86
WIAF-1796	5	71.59	WIAF-2177	6	50.57
WIAF-1869	5	83.99	WIAF-2106	6	51.04
WIAF-1721	5	94.22	WIAF-2810	6	59.2
WIAF-205	5	95.04	WIAF-538	6	72.55
WIAF-2941	5	98.58	WIAF-4584	6	77.27
WIAF-512	5	99.89	WIAF-3700	6	79.56
WIAF-3553	5	100.29	WIAF-3444	6	82.12
WIAF-1594	5	106.6	WIAF-3843	6	85.55
WIAF-3198	5	107.18	WIAF-2690	6	87.39
WIAF-3199	5	107.18	WIAF-3097	6	88.06
WIAF-4558	5	111.13	WIAF-2877	6	91.42
WIAF-3339	5	112.7	WIAF-3161	6	94.14
WIAF-2723	5	118.43	WIAF-3158	6	94.16
WIAF-3210	5	128.24	WIAF-922	6	94.83
WIAF-1804	5	128.4	WIAF-4572	6	95.35
WIAF-287	5	129.3	WIAF-1670	6	95.54
WIAF-3133	5	130.14	WIAF-1541	6	96.83
WIAF-2108	5	132.9	WIAF-3799	6	97.3
WIAF-3547	5	135.31	WIAF-3746	6	97.51
WIAF-3246	5	136.75	WIAF-2956	6	102.94
WIAF-1331	5	140.75	WIAF-1476	6	109.88
WIAF-3559	5	146.07	WIAF-577	6	109.99
WIAF-209	5	152.09	WIAF-3344	6	111.27
WIAF-3402	5	154.16	WIAF-3245	6	117.81
WIAF-2087	5	169.7	WIAF-207	6	118.98
WIAF-3788	5	171.16	WIAF-3877	6	120.35
WIAF-1827	5	171.33	WIAF-985	6	127.12
WIAF-2023	5	173.82	WIAF-224	6	128.09
WIAF-3137	5	204.79	WIAF-3191	6	128.5
WIAF-3000	5	205.22	WIAF-3723	6	128.63
WIAF-2627	6	1.25	WIAF-3148	6	128.65
WIAF-3613	6	1.36	WIAF-726	6	128.65
WIAF-2709	6	2.41	WIAF-584	6	128.69

续表

SNP名称	染色体位置	遗传距离/cM	SNP名称	染色体位置	遗传距离/cM
WIAF-3081	6	129.72	WIAF-3187	7	135.26
WIAF-736	6	132.47	WIAF-1680	7	135.8
WIAF-737	6	132.48	WIAF-2364	7	137.25
WIAF-3531	6	132.85	WIAF-983	7	137.71
WIAF-1674	6	134.75	WIAF-671	7	139.85
WIAF-3474	6	141.01	WIAF-672	7	139.85
WIAF-3645	6	148.62	WIAF-2152	7	140.33
WIAF-3312	6	153.86	WIAF-3292	7	144.25
WIAF-2767	6	153.88	WIAF-3585	7	145.88
WIAF-592	6	162.59	WIAF-3285	7	145.92
WIAF-3676	6	164.85	WIAF-4573	7	150.32
WIAF-3677	6	164.91	WIAF-4518	7	151.36
WIAF-3806	6	172.8	WIAF-3650	7	155.99
WIAF-1955	6	176.79	WIAF-2651	7	156.67
WIAF-762	6	179.07	WIAF-2287	7	181.84
WIAF-982	6	179.16	WIAF-2264	7	181.85
WIAF-3452	6	180.76	WIAF-1627	7	184.82
WIAF-915	6	192.1	WIAF-3238	7	195.8
WIAF-3328	6	192.43	WIAF-1785	7	199.25
WIAF-3355	6	193.96	WIAF-1991	7	199.53
WIAF-3704	6	203.58	WIAF-3212	7	200.3
WIAF-3354	6	230.11	WIAF-2817	7	201.07
WIAF-3064	7	0	WIAF-2815	7	202
WIAF-350	7	8.12	WIAF-1977	7	215.34
WIAF-2601	7	8.43	WIAF-3834	8	3.61
WIAF-3390	7	21.69	WIAF-1992	8	9.51
WIAF-4506	7	25.7	WIAF-283	8	12.21
WIAF-4539	7	25.89	WIAF-2927	8	12.9
WIAF-514	7	55.89	WIAF-1897	8	13.4
WIAF-263	7	55.9	WIAF-1898	8	13.4
WIAF-4504	7	60.54	WIAF-3345	8	15.69
WIAF-2616	7	60.57	WIAF-3346	8	15.72
WIAF-3123	7	65.85	WIAF-1085	8	24.81
WIAF-678	7	65.89	WIAF-1745	8	24.83
WIAF-902	7	68.5	WIAF-496	8	32.68
WIAF-827	7	76.47	WIAF-2120	8	38.13
WIAF-828	7	76.7	WIAF-865	8	40.63
WIAF-1810	7	82.03	WIAF-1517	8	44.31
WIAF-636	7	83.18	WIAF-3157	8	44.47
WIAF-4576	7	83.19	WIAF-3156	8	44.48
WIAF-3321	7	101.5	WIAF-414	8	46.25
WIAF-806	7	112.47	WIAF-3420	8	46.28
WIAF-1218	7	119.48	WIAF-357	8	46.42
WIAF-4521	7	121.01	WIAF-996	8	46.73
WIAF-4537	7	121.04	WIAF-2809	8	48.14
WIAF-3603	7	124.82	WIAF-867	8	48.15
WIAF-3422	7	126.04	WIAF-819	8	49
WIAF-1772	7	127.21	WIAF-2525	8	50.3
WIAF-4117	7	129.95	WIAF-832	8	51.56
WIAF-2086	7	130.26	WIAF-2737	8	54
WIAF-807	7	132.59	WIAF-2738	8	54.02

续表

SNP名称	染色体位置	遗传距离/cM	SNP名称	染色体位置	遗传距离/cM
WIAF-716	8	56.54	WIAF-3206	9	23.75
WIAF-2980	8	60.11	WIAF-2945	9	26.33
WIAF-370	8	60.28	WIAF-3688	9	26.54
WIAF-3218	8	65.62	WIAF-3600	9	27.71
WIAF-1693	8	67.76	WIAF-3093	9	44.7
WIAF-3573	8	68.18	WIAF-459	9	44.77
WIAF-3054	8	69.3	WIAF-2759	9	46.97
WIAF-2374	8	69.81	WIAF-1919	9	48.24
WIAF-3621	8	69.82	WIAF-1899	9	49.39
WIAF-2377	8	70.91	WIAF-1900	9	49.42
WIAF-472	8	72.47	WIAF-2372	9	56.14
WIAF-355	8	72.84	WIAF-4131	9	56.88
WIAF-1482	8	76.96	WIAF-4569	9	58.69
WIAF-2925	8	79.1	WIAF-3323	9	58.93
WIAF-1565	8	84.07	WIAF-975	9	61.82
WIAF-1580	8	89.27	WIAF-4540	9	64.81
WIAF-1885	8	90.82	WIAF-1634	9	64.89
WIAF-2084	8	91.39	WIAF-1450	9	68.06
WIAF-4516	8	98.32	WIAF-2782	9	69.99
WIAF-4509	8	98.35	WIAF-2783	9	69.99
WIAF-2268	8	99.23	WIAF-3813	9	70.8
WIAF-1846	8	100.28	WIAF-2165	9	86.46
WIAF-3341	8	102.27	WIAF-2133	9	90.92
WIAF-2312	8	104.31	WIAF-2275	9	91.58
WIAF-1320	8	104.33	WIAF-2567	9	94.45
WIAF-4536	8	104.34	WIAF-3520	9	94.64
WIAF-490	8	104.37	WIAF-3518	9	94.74
WIAF-2633	8	110.2	WIAF-960	9	94.74
WIAF-2226	8	113.99	WIAF-3831	9	96.74
WIAF-4035	8	114.47	WIAF-4544	9	104.33
WIAF-315	8	114.6	WIAF-3928	9	105.76
WIAF-3253	8	121.4	WIAF-1764	9	106.56
WIAF-341	8	130.9	WIAF-433	9	107.15
WIAF-3681	8	137.33	WIAF-3554	9	110.29
WIAF-761	8	137.79	WIAF-4508	9	111.39
WIAF-3231	8	143.76	WIAF-1412	9	114.09
WIAF-989	8	147.66	WIAF-4523	9	114.09
WIAF-1516	8	155.54	WIAF-2155	9	128.42
WIAF-583	8	155.82	WIAF-3477	9	129.27
WIAF-3033	8	163.14	WIAF-3358	9	129.84
WIAF-4524	8	164.32	WIAF-233	9	129.92
WIAF-1584	8	166.68	WIAF-951	9	130.48
WIAF-1027	8	168.85	WIAF-3036	9	136.95
WIAF-1186	8	169.13	WIAF-3801	9	139.77
WIAF-2402	8	170.79	WIAF-2953	9	140.48
WIAF-1029	8	178.66	WIAF-2721	9	140.6
WIAF-3881	9	0.34	WIAF-3938	9	141.15
WIAF-988	9	10.02	WIAF-3472	9	141.33
WIAF-3527	9	14.17	WIAF-2950	9	142.35
WIAF-2714	9	19.16	WIAF-2095	9	148.41
WIAF-1582	9	19.9	WIAF-2094	9	148.42

续表

SNP名称	染色体位置	遗传距离/cM	SNP名称	染色体位置	遗传距离/cM
WIAF-3153	9	156.98	WIAF-3019	11	18.77
WIAF-3035	9	157.02	WIAF-3129	11	18.86
WIAF-3025	9	157.08	WIAF-3125	11	18.9
WIAF-3310	9	157.08	WIAF-3139	11	18.9
WIAF-3736	9	164.96	WIAF-4592	11	19.89
WIAF-3006	9	166.96	WIAF-3023	11	25.18
WIAF-1742	10	10.09	WIAF-4530	11	37.24
WIAF-2581	10	15.49	WIAF-2046	11	40.22
WIAF-3537	10	16.5	WIAF-2311	11	40.27
WIAF-1511	10	27.68	WIAF-588	11	43.69
WIAF-272	10	31.56	WIAF-1949	11	43.98
WIAF-3614	10	42.47	WIAF-1307	11	48.25
WIAF-3908	10	45.56	WIAF-3007	11	49.68
WIAF-966	10	61.17	WIAF-2983	11	66.6
WIAF-2158	10	65.17	WIAF-2768	11	66.68
WIAF-3281	10	67.18	WIAF-2769	11	66.68
WIAF-4055	10	70.34	WIAF-2822	11	67.19
WIAF-3797	10	70.39	WIAF-2823	11	67.19
WIAF-1145	10	70.42	WIAF-4525	11	74.78
WIAF-3058	10	71.81	WIAF-2220	11	76.36
WIAF-2729	10	71.92	WIAF-1862	11	77.28
WIAF-3130	10	72.28	WIAF-1059	11	78.73
WIAF-3136	10	72.28	WIAF-3659	11	78.79
WIAF-303	10	76.37	WIAF-1529	11	79.13
WIAF-3110	10	79.84	WIAF-1866	11	79.84
WIAF-3899	10	88.16	WIAF-381	11	80.75
WIAF-3637	10	94.3	WIAF-4557	11	80.84
WIAF-3711	10	96.06	WIAF-380	11	80.87
WIAF-310	10	96.91	WIAF-2195	11	81.01
WIAF-308	10	97.37	WIAF-228	11	85.13
WIAF-744	10	98.09	WIAF-842	11	89
WIAF-2805	10	102.28	WIAF-2569	11	95.09
WIAF-1799	10	103.56	WIAF-2897	11	102.07
WIAF-1585	10	103.77	WIAF-2313	11	102.23
WIAF-2438	10	104.9	WIAF-3203	11	103.01
WIAF-4519	10	108.83	WIAF-3071	11	103.81
WIAF-3691	10	129.98	WIAF-1564	11	105.35
WIAF-2303	10	131.44	WIAF-3059	11	109.71
WIAF-2588	10	132.26	WIAF-3055	11	109.73
WIAF-2911	10	137.24	WIAF-1950	11	110.58
WIAF-4070	10	138.88	WIAF-1542	11	112.64
WIAF-901	10	144.51	WIAF-3012	11	119.75
WIAF-271	10	151.85	WIAF-1525	11	121.02
WIAF-1676	10	152.45	WIAF-3653	11	124.86
WIAF-3761	10	156.49	WIAF-2758	11	124.88
WIAF-3299	10	156.6	WIAF-1588	11	126.3
WIAF-3213	10	164.91	WIAF-3186	11	127.96
WIAF-3630	10	179.07	WIAF-3770	11	129
WIAF-991	11	16	WIAF-3769	11	129.01
WIAF-3297	11	16.23	WIAF-1247	11	130.7
WIAF-2043	11	16.47	WIAF-3882	11	131.06

SNP名称	染色体位置	遗传距离/cM	SNP名称	染色体位置	遗传距离/cM
WIAF-1694	11	132.16	WIAF-1520	12	129.46
WIAF-712	11	134.86	WIAF-3911	12	132.8
WIAF-713	11	134.95	WIAF-1918	12	136.13
WIAF-3291	11	135.29	WIAF-2467	12	154.19
WIAF-971	11	141.63	WIAF-3624	12	154.59
WIAF-432	11	142.66	WIAF-3594	13	11.14
WIAF-3838	11	143.48	WIAF-2385	13	12.09
WIAF-3745	11	147.24	WIAF-973	13	12.15
WIAF-1463	11	161	WIAF-3070	13	15.6
WIAF-1767	11	161	WIAF-4575	13	19.03
WIAF-2696	11	161	WIAF-2781	13	36.92
WIAF-3053	11	161	WIAF-2138	13	42.48
WIAF-3049	12	18.22	WIAF-3367	13	44.52
WIAF-2689	12	21.69	WIAF-1544	13	47.86
WIAF-1430	12	21.79	WIAF-1353	13	47.88
WIAF-1805	12	22.02	WIAF-2713	13	50.28
WIAF-3906	12	31.77	WIAF-3543	13	50.99
WIAF-4548	12	32.95	WIAF-1858	13	51.57
WIAF-879	12	35.59	WIAF-505	13	51.78
WIAF-3296	12	44.8	WIAF-3384	13	70.17
WIAF-2679	12	50.83	WIAF-2754	13	70.31
WIAF-4045	12	55.17	WIAF-3084	13	71.41
WIAF-641	12	55.51	WIAF-1318	13	72.63
WIAF-4526	12	57.04	WIAF-2207	13	77.24
WIAF-3915	12	57.94	WIAF-3756	13	94.06
WIAF-2363	12	59.7	WIAF-2706	13	121.67
WIAF-1988	12	63.03	WIAF-2555	13	161
WIAF-2787	12	63.98	WIAF-1026	14	0
WIAF-2785	12	64.14	WIAF-436	14	0
WIAF-533	12	69.3	WIAF-3887	14	3.41
WIAF-1980	12	70.53	WIAF-3227	14	7.57
WIAF-463	12	71.53	WIAF-4082	14	10.31
WIAF-2420	12	71.75	WIAF-2867	14	18.14
WIAF-2423	12	71.75	WIAF-2868	14	18.25
WIAF-2661	12	74.34	WIAF-729	14	19.02
WIAF-2658	12	74.35	WIAF-3249	14	37.74
WIAF-3738	12	83.07	WIAF-3383	14	40.64
WIAF-3140	12	83.18	WIAF-2789	14	40.97
WIAF-818	12	85.81	WIAF-2919	14	43.31
WIAF-3644	12	86.93	WIAF-4542	14	43.59
WIAF-3330	12	87.52	WIAF-694	14	45.71
WIAF-1643	12	88.98	WIAF-1997	14	48.12
WIAF-1053	12	104.84	WIAF-3078	14	65.07
WIAF-377	12	105.33	WIAF-3777	14	69.24
WIAF-3063	12	105.54	WIAF-3713	14	69.46
WIAF-3549	12	110.7	WIAF-402	14	70.49
WIAF-4545	12	113.53	WIAF-1778	14	71.08
WIAF-3302	12	124.31	WIAF-1036	14	72.47
WIAF-552	12	124.7	WIAF-823	14	72.99
WIAF-2718	12	129.31	WIAF-295	14	73
WIAF-2716	12	129.42	WIAF-294	14	73.06

续表

SNP名称	染色体位置	遗传距离/cM	SNP名称	染色体位置	遗传距离/cM
WIAF-2677	14	73.23	WIAF-1945	15	82.62
WIAF-246	14	76.73	WIAF-2012	15	85.37
WIAF-3500	14	80.65	WIAF-1490	15	90.11
WIAF-478	14	85.18	WIAF-2489	15	94.47
WIAF-2735	14	85.2	WIAF-4547	15	94.54
WIAF-3237	14	85.62	WIAF-3876	15	97.14
WIAF-4259	14	95.72	WIAF-3683	15	97.37
WIAF-835	14	108.95	WIAF-2321	15	97.77
WIAF-1602	14	110.53	WIAF-2630	15	97.88
WIAF-646	14	110.53	WIAF-3106	15	98.11
WIAF-2218	14	111.62	WIAF-3111	15	98.12
WIAF-2233	14	112.18	WIAF-3114	15	98.12
WIAF-2301	14	122.02	WIAF-3022	15	98.41
WIAF-2302	14	122.04	WIAF-3699	15	103.31
WIAF-874	14	124.91	WIAF-444	15	105.08
WIAF-3184	14	132.99	WIAF-2849	15	117.87
WIAF-3185	14	132.99	WIAF-2931	15	161
WIAF-1895	14	134.41	WIAF-1514	16	7.68
WIAF-2041	14	161	WIAF-627	16	12
WIAF-4531	14	161	WIAF-924	16	12
WIAF-4513	15	5.77	WIAF-3828	16	16.36
WIAF-4593	15	5.77	WIAF-510	16	27.76
WIAF-4595	15	7.41	WIAF-1614	16	27.96
WIAF-4597	15	7.77	WIAF-3705	16	27.96
WIAF-4551	15	20.87	WIAF-1909	16	29.48
WIAF-4590	15	27.31	WIAF-2343	16	29.54
WIAF-802	15	27.32	WIAF-1156	16	43.32
WIAF-2802	15	28.71	WIAF-2652	16	44.07
WIAF-1283	15	32.13	WIAF-3574	16	46.49
WIAF-791	15	33.37	WIAF-1870	16	46.98
WIAF-2889	15	36.75	WIAF-606	16	49.9
WIAF-2887	15	39.42	WIAF-2136	16	58.74
WIAF-3863	15	42.99	WIAF-1971	16	68.82
WIAF-3907	15	43.02	WIAF-2618	16	68.92
WIAF-317	15	43.22	WIAF-1697	16	75.43
WIAF-618	15	44.65	WIAF-889	16	83.68
WIAF-2949	15	46.11	WIAF-2752	16	85.96
WIAF-2227	15	46.42	WIAF-2750	16	85.98
WIAF-1861	15	46.66	WIAF-2997	16	90.35
WIAF-336	15	51.19	WIAF-3421	16	92
WIAF-88	15	52.39	WIAF-3425	16	92
WIAF-3287	15	52.43	WIAF-3855	16	92
WIAF-3890	15	54.74	WIAF-3714	16	92.35
WIAF-1635	15	56.58	WIAF-1812	16	97.74
WIAF-4515	15	61.97	WIAF-2399	16	99.18
WIAF-534	15	64.15	WIAF-1474	16	99.23
WIAF-425	15	65.45	WIAF-3504	16	99.38
WIAF-1000	15	65.84	WIAF-3494	16	106.59
WIAF-362	15	68.25	WIAF-3270	16	106.94
WIAF-2576	15	68.46	WIAF-1139	16	107.59
WIAF-3619	15	69.03	WIAF-2650	16	110.63

SNP名称	染色体位置	遗传距离/cM	SNP名称	染色体位置	遗传距离/cM
WIAF-673	16	111.13	WIAF-2858	17	116.98
WIAF-1628	16	161	WIAF-3305	17	121.36
WIAF-2142	16	161	WIAF-1505	18	4.04
WIAF-2860	16	161	WIAF-1626	18	12.3
WIAF-2861	16	161	WIAF-1001	18	32.16
WIAF-3068	16	161	WIAF-1002	18	32.32
WIAF-974	16	161	WIAF-2971	18	41.64
WIAF-408	17	0	WIAF-2824	18	41.66
WIAF-409	17	0	WIAF-2137	18	43.18
WIAF-821	17	9.84	WIAF-2780	18	50.45
WIAF-918	17	14.07	WIAF-1468	18	61.19
WIAF-39	17	14.12	WIAF-3204	18	61.71
WIAF-3660	17	17.52	WIAF-907	18	61.84
WIAF-4585	17	18.25	WIAF-4147	18	66.09
WIAF-748	17	28.86	WIAF-1811	18	66.99
WIAF-747	17	29.15	WIAF-3013	18	67.69
WIAF-1274	17	34.78	WIAF-1775	18	74
WIAF-3197	17	41.14	WIAF-700	18	92.74
WIAF-2573	17	43.43	WIAF-210	18	93.21
WIAF-3889	17	43.57	WIAF-4080	18	97.49
WIAF-1519	17	52.83	WIAF-3180	18	99.36
WIAF-2293	17	55.65	WIAF-331	18	101.74
WIAF-3030	17	58.17	WIAF-801	18	115.85
WIAF-908	17	61.54	WIAF-693	18	119.26
WIAF-1996	17	61.85	WIAF-1119	18	123.8
WIAF-423	17	64.85	WIAF-2351	18	161
WIAF-422	17	64.92	WIAF-923	19	15.27
WIAF-3051	17	65.14	WIAF-894	19	21.73
WIAF-1741	17	67.04	WIAF-893	19	21.79
WIAF-1598	17	67.73	WIAF-3393	19	32.89
WIAF-752	17	68.02	WIAF-3706	19	39.09
WIAF-4164	17	70.3	WIAF-118	19	41.19
WIAF-1164	17	71.18	WIAF-608	19	47.96
WIAF-4554	17	71.3	WIAF-2857	19	49.63
WIAF-1133	17	73.82	WIAF-3276	19	51.57
WIAF-1134	17	73.86	WIAF-839	19	52.82
WIAF-1021	17	74.25	WIAF-1533	19	52.96
WIAF-3780	17	74.29	WIAF-2574	19	59.66
WIAF-3779	17	74.3	WIAF-1948	19	60.55
WIAF-507	17	76.22	WIAF-786	19	61.06
WIAF-1652	17	76.28	WIAF-2206	19	62.58
WIAF-2874	17	81.74	WIAF-2641	19	69.21
WIAF-511	17	86.09	WIAF-978	19	69.3
WIAF-1138	17	93.4	WIAF-2063	19	69.68
WIAF-1325	17	94.5	WIAF-2064	19	69.68
WIAF-2375	17	95.11	WIAF-3873	19	69.84
WIAF-2760	17	101.89	WIAF-4510	19	73.34
WIAF-2560	17	107.06	WIAF-3546	19	75.2
WIAF-2450	17	112.69	WIAF-3607	19	88.94
WIAF-2445	17	113.15	WIAF-2639	19	95.43
WIAF-2413	17	116.26	WIAF-2637	19	97.8

503

续表

SNP名称	染色体位置	遗传距离/cM	SNP名称	染色体位置	遗传距离/cM
WIAF-2101	19	103.18	WIAF-624	21	34.81
WIAF-4517	19	104.93	WIAF-1500	21	35.02
WIAF-3024	19	105.19	WIAF-3496	21	35.39
WIAF-1488	19	105.71	WIAF-1943	21	37.13
WIAF-2662	19	109.51	WIAF-2477	21	50.92
WIAF-2129	20	0.91	WIAF-2436	21	161
WIAF-2119	20	8.9	WIAF-4514	21	161
WIAF-730	20	17.76	WIAF-3618	22	0
WIAF-3220	20	28.24	WIAF-843	22	2.79
WIAF-3221	20	28.56	WIAF-2544	22	5.54
WIAF-3511	20	30.78	WIAF-3824	22	9
WIAF-3612	20	32.93	WIAF-1706	22	21.16
WIAF-3041	20	43.93	WIAF-2057	22	22.83
WIAF-3941	20	44.38	WIAF-3880	22	25.23
WIAF-2128	20	44.4	WIAF-2538	22	36.66
WIAF-3927	20	53.91	WIAF-230	22	37.28
WIAF-2019	20	60.03	WIAF-1974	22	37.91
WIAF-266	20	68.28	WIAF-1672	22	38.19
WIAF-1648	20	70.31	WIAF-1673	22	38.25
WIAF-1853	20	75.9	WIAF-3864	22	38.94
WIAF-2963	20	80.82	WIAF-1266	22	44.06
WIAF-1969	20	94.08	WIAF-1816	22	50.86
WIAF-3497	21	8.08	WIAF-1904	22	50.89
WIAF-3498	21	8.36	WIAF-3065	22	161
WIAF-599	21	17.53	WIAF-3067	22	161
WIAF-3562	21	18.91	WIAF-2669	X	18.97
WIAF-1682	21	23.86	WIAF-4550	X	73.33
WIAF-559	21	24.14	WIAF-1290	X	111.15
WIAF-4546	21	26.04	WIAF-1872	X	124.12
WIAF-3508	21	32.05	WIAF-3737	X	124.19
WIAF-2643	21	33.99	WIAF-2589	X	160.92

504

参考文献

Ahrendt S.A., Halachmi S., Chow J.T., Wu L., Halachmi N., Yang S.C., Wehage S., Jen J., and Sidransky D. 1999. Rapid p53 sequence analysis in primary lung cancer using an oligonucleotide probe array. *Proc. Natl. Acad. Sci.* **96:** 7382–7387.

Albertson D.G., Ylstra B., Segraves R., Collins C., Dairkee S.H., Kowbel D., Kuo W.L., Gray J.W., and Pinkel D. 2000. Quantitative mapping of amplicon structure by array CGH identifies CYP24 as a candidate oncogene. *Nat. Genet.* **25:** 144–146.

Alizadeh A.A., Eisen M.B., Davis R.E., Ma C., Lossos I.S., Rosenwald A., Boldrick J.C., Sabet H., Tran T., Yu X., Powell J.I., Yang L., Marti G.E., Moore T., Hudson J. Jr., Lu L., Lewis D.B., Tibshirani R., Sherlock G., Chan W.C., Greiner T.C., Weisenburger D.D., Armitage J.O., Warnke R., Staudt L.M., et al. 2000. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* **403:** 503–511.

Boyd K.E. and Farnham P.J. 1999. Coexamination of site-specific transcription factor binding and promoter activity in living cells. *Mol. Cell. Biol.* **19:** 8393–8399.

Chakravarti A. 1998. It's raining SNPs, hallelujah? *Nat. Genet.* **19:** 216–217.

Chee M., Yang R., Hubbell E., Berno A., Huang X.C., Stern D., Winkler J., Lockhart D.J., Morris M.S., and Fodor S.P. 1996. Accessing genetic information with high-density DNA arrays. *Science* **274:** 610–614.

Chen J., Iannone M.A., Li M.S., Taylor J.D., Rivers P., Nelsen A.J., Slentz-Kesler K.A., Roses A., and Weiner M.P. 2000. A microsphere-based assay for multiplexed single nucleotide polymorphism analysis using single base chain extension. *Genome Res.* **10:** 549–557.

Chen X. and Kwok P.Y. 1999. Homogeneous genotyping assays for single nucleotide polymorphisms with fluorescence resonance energy transfer detection. *Genet. Anal.* **14:** 157–163.

Chen X., Levine L., and Kwok P.Y. 1999. Fluorescence polarization in homogeneous nucleic acid analysis. *Genome Res.* **9:** 492–498.

Cheung V.G., Nowak N., Jang W., Kirsch I.R., Zhao S., Chen X.N., Furey T.S., Kim U.J., Kuo WL, Olivier M., Conroy J., Kasprzyk A., Massa H., Yonescu R., Sait S., Thoreen C., Snijders A.. Lemyre E., Bailey J.A., Bruzel A., Burrill W.D., Clegg S.M., Collins S., Dhami P., Friedman C, Han C.S., Herrick S., Lee J., Ligon A.H., Lowry S., Morley M., Narasimhan S., Osoegawa K., Peng Z., Plajzer-Frick I., Quade B.J., Scott D., Sirotkin K., Thorpe A.A., Gray J.W., Hudson J, Pinkel D., Ried T., Rowen L., Shen-Ong G.L., Strausberg R.L., Birney E., Callen D.F., Cheng J.F., Cox D.R., Doggett N.A., Carter N.P., Eichler E.E., Haussler D., Korenberg J.R., Morton C.C., Albertson D., Schuler G., de Jong P.J., Trask B.J.; The BAC Resource Consortium. 2001. Integration of cytogenetic landmarks into the draft sequence of the human genome. *Nature* **409:** 953–958.

Collins F.S., Patrinos A., Jordan E., Chakravarti A., Gesteland R., and Walters L. 1998. New goals for the U.S. Human Genome Project: 1998–2003. *Science* **282:** 682–689.

Cutler D.J., Zwick M.E., Carrasquillo M.M., Yohn C.T., Tobin K.P., Kashuk C., Mathews D.J., Shah N.A., Eichler E.E., Warrington J.A., and Chrakravarti A. 2001. High-throughput variation detection and genotyping using microarrays. *Genome Res.* **11:** 1913–1925.

Deloukas P., Schuler G.D., Gyapay G., Beasley E.M., Soderlund C., Rodriguez-Tome P., Hui L., Matise T.C., McKusick K.B., Beckmann J.S., Bentolila S., Bihoreau M., Birren B.B., Browne J., Butler A., Castle A.B., Chiannilkulchai N., Clee C., Day P.J., Dehejia A., Dibling T., Drouot N., Duprat S., Fizames C., Bentley D.R., et al. 1998. A physical map of 30,000 human genes. *Science* **282:** 744–746.

DeRisi J., Penland L., Brown P.O., Bittner M.L., Meltzer P.S., Ray M., Chen Y., Su Y.A., Trent J.M., et al. 1996. Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat. Genet.* **14:** 457–460.

du Manoir S., Speicher M.R., Joos S., Schrock E., Popp S., Dohner H., Kovacs G., Robert-Nicoud M., Lichter P., and Cremer T. 1993. Detection of complete and partial chromosome gains and losses by comparative genomic in situ hybridization. *Hum. Genet.* **90:** 590–610.

Dracopoli N.C. and Fogh J. 1983. Loss of heterozygosity in cultured human tumor cell lines. *J. Natl. Cancer Inst.* **70:** 83–87.

Eggerding F.A. 1995. A one-step coupled amplification and oligonucleotide ligation procedure for multiplex genetic typing. *PCR Methods Appl.* **4:** 337–345.

Fan J.B., Chen X., Halushka M.K., Berno A., Huang X., Ryder T., Lipshutz R.J., Lockhart D.J., and Chakravarti A. 2000. Parallel genotyping of human SNPs using generic high-density oligonucleotide tag arrays. *Genome Res.* **10:** 853–860.

Fortina P., Conant R., Monokian G., Dotti G., Parrella T., Hitchcock W., Kant J., Scanlin T., Rappaport E. Schwartz E., et al. 1992. Non-radioactive detection of the most common mutations in the cystic fibrosis transmembrane conductance regulator gene by multiplex allele-specific polymerase chain reaction. *Hum. Genet.* **90:** 375–378.

Gall K., Pavelic J., Fadro-Santel D., Poljak M., and Pavelic K. 1993. DNA amplification by polymerase chain reaction from brain tissues embedded in paraffin. *Int. J. Exp. Pathol.* **74:** 333–337.

Geschwind D.H., Gregg J., Boone K., Karrim J., Pawlikowska-Haddal A., Rao E., Ellison J., Ciccodicola A., D'Urso M., Woods R., Rappold G.A., Swerdloff R., and Nelson S.F. 1998. Klinefelter's syndrome as a model of anomalous cerebral laterality: Testing gene dosage in the X chromosome pseudoautosomal region using a DNA microarray. *Dev. Genet.* **23:** 215–229.

Greer C.E., Wheeler C.M., and Manos M.M. 1995. PCR amplification from paraffin-embedded tissues: Sample preparation and the effects of fixation. In *PCR primer: A laboratory manual* (ed. C.W. Dieffenbach and G.S. Dveksler), pp. 99–112. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Hacia J.G. and Collins F.S. 1999. Mutational analysis using oligonucleotide microarrays. *J. Med. Genet.* **36:** 730–736.

Hacia J.G., Brody L.C., Chee M.S., Fodor S.P., and Collins F.S. 1996. Detection of heterozygous mutations in BRCA1 using high density oligonucleotide arrays and two-color fluorescence analysis. *Nat. Genet.* **14:** 441–447.

Haff L.A. and Smirnov I.P. 1997. Single-nucleotide polymorphism identification assays using a thermostable DNA polymerase and delayed extraction MALDI-TOF mass spectrometry. *Genome Res.* **7:** 378–388.

Hemminki A., Tomlinson I., Markie D., Jarvinen H., Sistonen P., Bjorkqvist A.M., Knuutila S., Salovaara R., Bodmer W., Shibata D., de la Chapelle A., and Aaltonen L.A. 1997. Localization of a susceptibility locus for Peutz-Jeghers syndrome to 19p using comparative genomic hybridization and targeted linkage analysis. *Nat. Genet.* **15:** 87–90.

Hirschhorn J.N., Sklar P., Lindblad-Toh K., Lim Y.M., Ruiz-Gutierrez M., Bolk S., Langhorst B., Schaffner S., Winchester E., and Lander E.S. 2000. SBE-TAGS: An array-based method for efficient single-nucleotide polymorphism genotyping. *Proc. Natl. Acad. Sci.* **97:** 12164–12169.

Hodgson G., Hager J.H., Volik S., Hariono S., Wernick M., Moore D., Albertson D.G., Pinkel D., Collins C., Hanahan D., and Gray J.W. 2001. Genome scanning with array CGH delineates regional alterations in mouse islet carcinomas. *Nat. Genet.* **29:** 459–464.

Iyer V.R., Horak C.E., Scafe C.S., Botstein D., Snyder M., and Brown P.O. 2001. Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature* **409:** 533–538.

Jain A.N., Tokuyasu T.A., Snijders A.M., Segraves R., Albertson D.G., and Pinkel D. 2002. Fully automatic quantification of microarray image data. *Genome Res.* **12:** 325–332.

Kallioniemi A., Kallioniemi O.P., Sudar D., Rutovitz D., Gray J.W., Waldman F., and Pinkel D. 1992. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science* **258:** 818–821.

———. 1993. Comparative genomic hybridization: A rapid new method for detecting and mapping DNA amplification in tumors. *Semin. Cancer Biol.* **4:** 41–46.

Kiko H., Niggemann E., and Ruger W. 1979. Physical mapping of the restriction fragments obtained from bacteriophage T4 dC-DNA with the restriction endonucleases *Sma*I, *Kpn*I and *Bgl*II. *Mol. Gen. Genet.* **172:** 303–312.

Klein C.A., Schmidt-Kittler O., Schardt J.A., Pantel K., Speicher M.R., and Riethmuller G. 1999. Comparative genomic hybridization, loss of heterozygosity, and DNA sequence analysis of single cells. *Proc. Natl. Acad. Sci.* **96:** 4494–4499.

Kozal M.J., Shah N., Shen N., Yang R., Fucini R., Merigan T.C., Richman D.D., Morris D., Hubbell E., Chee M., and Gingeras T.R. 1996. Extensive polymorphisms observed in HIV-1 clade B protease gene using high-density oligonucleotide arrays. *Nat. Med.* **2:** 753–759.

Kwok P.Y. 2000. High-throughput genotyping assay approaches. *Pharmacogenomics* **1:** 95–100.

Lander E.S., Linton L.M., Birren B., Nusbaum C., Zody M.C., Baldwin J., Devon K., Dewar K., Doyle M., FitzHugh W., et al. 2001. Initial sequencing and analysis of the human genome. *Nature* **409:** 860–921.

Lieb J.D., Liu X., Botstein D., and Brown P.O. 2001. Promoter-specific binding of Rap1 revealed by genome-wide maps of protein-DNA association. *Nat. Genet.* **28:** 327–334.

Lindblad-Toh K., Tanenbaum D.M., Daly M.J., Winchester E., Liu W.O., Villapakkam A., Stanton S.E., Larsson C., Hudson T.J., Johnson B.E., Lander E.S., and Meyerson M. 2000a. Loss-of-heterozygosity analysis of small cell lung carcinoma using single nucleotide polymorphism arrays. *Nat. Biotechnol.* **18:** 1001–1005.

Lindblad-Toh K., Winchester E., Daly M.J., Wang D.G., Hirschhorn J.N., Laviolette J.P., Ardlie K., Reich D.E., Robinson E., Sklar P., Shah N., Thomas D., Fan J.B., Gingeras T., Warrington J., Patil N., Hudson T.J., and Lander E.S. 2000b. Large-scale discovery and genotyping of single-nucleotide polymorphisms in the mouse. *Nat. Genet.* **24:** 381–386.

Lisitsyn N. and Wigler M. 1993. Cloning the differences between two complex genomes. *Science* **259:** 946–951.

Lisitsyn N.A., Rosenberg M.V., Launer G.A., Wagner L.L., Potapov V.K., Kolesnik T.B., and Sverdlov E.D. 1993. A method for isolation of sequences missing in one of two related genomes. *Mol. Gen. Mikrobiol. Virusol.* **3:** 26–29.

Lisitsyn N.A., Lisitsina N.M., Dalbagni G., Barker P., Sanchez C.A., Gnarra J., Linehan W.M., Reid B.J., and Wigler M.H. 1995. Comparative genomic analysis of tumors: Detection of DNA losses and amplification. *Proc. Natl. Acad. Sci.* **92:** 151–155.

Livak K.J., Flood S.J., Marmaro J., Giusti W., and Deetz K. 1995. Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. *PCR Methods Appl.* **4:** 357–362.

Lucito R., Nakimura M., West J.A., Han Y., Chin K., Jensen K., McCombie R., Gray J.W., and Wigler M. 1998. Genetic analysis using genomic representations. *Proc. Natl. Acad. Sci.* **95:** 4487–4492.

Lucito R., West J., Reiner A., Alexander J., Esposito D., Mishra B., Powers S., Norton L., and Wigler M. 2000. Detecting gene copy number fluctuations in tumor cells by microarray analysis of genomic representations. *Genome Res.* **10:** 1726–1736.

Marth G., Yeh R., Minton M., Donaldson R., Li Q., Duan S., Davenport R., Miller R.D., and Kwok P.Y. 2001. Single-nucleotide polymorphisms in the public domain: How useful are they? *Nat. Genet.* **27:** 371–372.

Mei R., Galipeau P.C., Prass C., Berno A., Ghandour G., Patil N., Wolff R.K., Chee M.S., Reid B.J., and Lockhart D.J. 2000. Genome-wide detection of allelic imbalance using human SNPs and high-density DNA arrays. *Genome Res.* **10:** 1126–1137.

Monni O., Joensuu H., Franssila K., and Knuutila S. 1996. DNA copy number changes in diffuse large B-cell lymphoma—Comparative genomic hybridization study. *Blood* **87:** 5269–5278.

Orlando V. and Paro R. 1993. Mapping polycomb-repressed domains in the bithorax complex using in vivo formaldehyde cross-linked chromatin. *Cell* **75:** 1187–1198.

Pastinen T., Raitio M., Lindroos K., Tainola P., Peltonen L., and Syvanen A.C. 2000. A system for specific, high-throughput genotyping by allele-specific primer extension on microarrays. *Genome Res.* **10:** 1031–1042.

Patil N., Berno A.J., Hinds D.A., Barrett W.A., Doshi J.M., Hacker C.R., Kautzer C.R., Lee D.H., Marjoribanks C., McDonough D.P., Nguyen B.T.N., Norris M.C., Sheehan J.B., Shen N., Stern D., Stokowski R.P., Thomas D.J., Trulson M.O., Vyas K.R., Frazer K.A., Fodor S.P.A., Cox D.R. 2001. Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. *Science* **294:** 1719–1723.

Perou C.M., Sorlie T., Eisen M.B., van de Rijn M., Jeffrey S.S., Rees C.A., Pollack J.R., Ross D.T., Johnsen H., Akslen L.A., Fluge O., Pergamenschikov A., Williams C., Zhu S.X., Lonning P.E., Borresen-Dale A.L., Brown P.O., and Botstein D. 2000. Molecular portraits of human breast tumours. *Nature* **406:** 747–752.

Pinkel D., Segraves R., Sudar D., Clark S., Poole I., Kowbel D., Collins C., Kuo W.L., Chen C., Zhai Y., Dairkee S.H., Ljung B.M., Gray J.W., and Albertson D.G. 1998. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat. Genet.* **20:** 207–211.

Pollack J.R., Perou C.M., Alizadeh A.A., Eisen M.B., Pergamenschikov A., Williams C.F., Jeffrey S.S., Botstein D., and Brown P.O. 1999. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat. Genet.* **23:** 41–46.

Prince J.A. Feuk L., Howell W.M., Jobs M., Emahazion T., Blennow K., and Brookes A.J. 2001. Robust and accurate single nucleotide polymorphism genotyping by dynamic allele-specific hybridization (DASH): Design criteria and assay validation. *Genome Res.* **11:** 152–162.

Reid J.L., Iyer V.R., Brown P.O., and Struhl K. 2000. Coordinate regulation of yeast ribosomal protein genes is associated with targeted recruitment of Esa1 histone acetylase. *Mol. Cell* **6:** 1297–1307.

Ren B., Robert F., Wyrick J.J., Aparicio O., Jennings E.G., Simon I., Zeitlinger J., Schreiber J., Hannett N., Kanin E., Volkert T.L., Wilson C.J., Bell S.P., and Young R.A. 2000. Genome-wide location and function of DNA binding proteins. *Science* **290:** 2306–2309.

Ronaghi M., Nygren M., Lundeberg J., and Nyren P. 1999. Analyses of secondary structures in DNA by pyrosequencing. *Anal. Biochem.* **267:** 65–71.

Ryan D., Nuccie B., and Arvan D. 1999. Non-PCR-dependent detection of the factor V Leiden mutation from genomic DNA using a homogeneous invader microtiter plate assay. *Mol. Diagn.* **4:** 135–144.

Sachidanandam R., Weissman D., Schmidt S.C., Kakol J.M., Stein L.D., Marth G., Sherry S., Mullikin J.C., Mortimore B.J., Willey D.L., Hunt S.E., Cole C.G., Coggill P.C., Rice C.M., Ning Z., Rogers J., Bentley D.R., Kwok P.Y., Mardis E.R, Yeh R.T., Schultz B., Cook L., Davenport R., Dante M., Fulton L., Hillier L., Waterston R.H., McPherson J.D., Gilman B., Schaffner S., Van Etten W.J., Reich D., Higgins J., Daly M.J., Blumenstiel B., Baldwin J., Stange-Thomann N., Zody M.C., Linton L., Lander E.S., and Atshuler D. 2001. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* **409:** 928–933.

Sambrook J. and Russell D. 2001. *Molecular cloning: A laboratory manual*, third edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sapolsky R.J., Hsie L., Berno A., Ghandour G., Mittmann M., and Fan J.B. 1999. High-throughput polymorphism screening and genotyping with high-density oligonucleotide arrays. *Genet. Anal.* **14:** 187–192.

Schrock E., du Manoir S., Veldman T., Schoell B., Wienberg J., Ferguson-Smith M.A., Ning Y., Ledbetter D.H., Bar-Am I., Soenksen D., Garini Y., and Ried T. 1996. Multicolor spectral karyotyping of human chromosomes. *Science* **273:** 494–497.

Schuler G.D. 1997. Pieces of the puzzle: Expressed sequence tags and the catalog of human genes. *J. Mol. Med.* **75:** 694–698.

Shi M.M. 2001. Enabling large-scale pharmacogenetic studies by high-throughput mutation detection and genotyping technologies. *Clin. Chem.* **47:** 164–172.

Snijders A.M., Segraves R., Blackwood S., Pinkel D., and Albertson D.G. 2002. BAC microarray-based comparative genomic hybridization. *Methods Mol. Biol.* (in press).

Snijders A.M., Nowak N., Segraves R., Blackwood S., Brown N., Conroy J., Hamilton G., Hindle A.K., Huey B., Kimura K., Law S., Myambo K., Palmer J., Ylstra B., Yue J.P., Gray J.W., Jain A.N., Pinkel D., and Albertson D.G. 2001. Assembly of microarrays for genome-wide measurement of DNA copy number. *Nat. Genet.* **29:** 263–264.

Solinas-Toldo S., Lampel S., Stilgenbauer S., Nickolenko J., Benner A., Dohner H., Cremer T., and Lichter P. 1997. Matrix-based comparative genomic hybridization: Biochips to screen for genomic imbalances. *Genes Chromosomes Cancer* **20:** 399–407.

Strahl-Bolsinger S., Hecht A., Luo K., and Grunstein M. 1997. *SIR2* and *SIR4* interactions differ in core and extended telomeric heterochromatin in yeast. *Genes Dev.* **11:** 83–93.

Syvanen A.C. 2001. Accessing genetic variation: Genotyping single nucleotide polymorphisms. *Nat. Rev. Genet.* **2:** 930–942.

Thiagalingam S., Laken S., Willson J.K., Markowitz S.D., Kinzler K.W., Vogelstein B., and Lengauer C. 2001. Mechanisms underlying losses of heterozygosity in human colorectal cancers. *Proc. Natl. Acad. Sci.* **98:** 2698–2702.

Tyagi S., Bratu D.P., and Kramer F.R. 1998. Multicolor molecular beacons for allele discrimination. *Nat. Biotechnol.* **16:** 49–53.

Visakorpi T., Kallioniemi A.H., Syvanen A.C., Hyytinen E.R., Karhu R., Tammela T., Isola J.J., and Kallioniemi O.P. 1995. Genetic changes in primary and recurrent prostate cancer by comparative genomic hybridization. *Cancer Res.* **55:** 342–347.

Visakorpi T., Hyytinen E., Koivisto P., Tanner M., Keinanen R., Palmberg C., Palotie A., Tammela T., Isola J., and Kallioniemi O.P. 1995. In vivo amplification of the androgen receptor gene and progression of human prostate cancer. *Nat. Genet.* **9:** 401–406.

Walter M., Spillett D., Thomas P., Weissenbach J., and Goodfellow P. 1994. A method for constructing radiation hybrid maps of whole genomes. *Nat. Genet.* **7:** 22–28.

Wang D.G., Fan J.-B., Siao C.J., Berno A., Young P., Sapolsky R., Ghandour G., Perkins N., Winchester E., Spencer J., Kruglyak L., Stein L., Hsie L., Topaloglou T., Hubell E., Robinson E., Mittman M., Morris M.S., Shen N., Kilburn D., Rioux J., Nusbaum C., Rozen S., Hudson T.J., Lipshutz R., Chee M., and Lander E.S. 1998. Large-scale identification, mapping and genotyping of single-nucleotide polymorphisms in the human genome. *Science* **280:** 1077–1082.

Weissenbach J., Gyapay G., Dib C., Vignal A., Morissette J., Millasseau P., Vaysseix G., and Lathrop M. 1992. A second-generation linkage map of the human genome. *Nature* **359:** 794–801.

Wickman F.P., Lu M.L., Thykjaer T., Olesen S.H., Andersen L.D., Cordon-Cardo C., and Orntoft T.F. 2000. Evaluation of the performance of a p53 sequencing microarray chip using 140 previously sequenced bladder tumor samples. *Clin. Chem.* **46:** 1555–1561.

因特网资源

http://cmgm.stanford.edu/pbrown/mguide/HybChamber.pdf Hybridization chamber drawings, P. Brown Web Site, Stanford University

http://gai.nci.nih.gov/CHLC/ The Cooperative Human Linkage Center (CHLC)

http://genome.ucsc.edu/ The UCSC Human Genome Project Working Draft

http://genome-www.stanford.edu/~rees/ CaryoScope, Christian A. Rees, Stanford University

http://p53.curie.fr. p53 Web Site at the Institute Curie

http://www.corning.com/CMT/ Corning Microarray Technology

http://www.genome.wi.mit.edu Whitehead Institute/MIT Center for Genome Research

http://www.ncbi.nlm.nih.gov/blast/Blast.cgi Protein-protein BLAST search

http://www.progenetix.net Progenetix-Comparative Genomic Hybridization (CGH) database

http://www-shgc.stanford.edu/RH/index.html The Stanford Human Genome Center RH Server

第7章 基因芯片生物信息学简介

509

导言

面临着“洪水般的数据”，微阵列界正在争相发展能够有效地进行分析、系统地存储和共享微阵列数据的技术。本章介绍了一些当前正在使用的、用于管理急剧增长的信息的方法。

510

设计

如果未采用组织化的方法来进行微阵列实验，代价将会很昂贵，而且要花费大量的金钱，因此对于微阵列实验而言，最优设计尤为重要。PART Ⅰ中讨论了实验设计的四个方面。

PART Ⅰ 微阵列表达实验的设计

本部分论述一般的设计问题，如可供利用的微阵列玻片、标记反应中需要的RNA的量以及待检测样品的数量。更为特殊的设计考虑来自于与玻片微阵列进行双色竞争杂交时所固有的比较特性（Schena et al. 1995，1996），特别是样本之间是否要进行相互的直接比较，或者与共同的参考做对比，或者是通过这两种方式的组合。更为深入的讨论集中于研究多因子相互作用时带来的实验设计问题，如使用细胞因子的组合来处理细胞。因为微阵列通常同时检测数千个基因，即使是少量的噪声也会导致大量的假阳性。

通过重复实验，包括染料交换，能够对结果进行平均处理，这可以提高准确性，并使得对数据质量和显著性的统计学评价变得容易（Lee et al. 2000；Kerr and Churchill 2001；Tseng et al. 2001）。最后，这一部分讨论了重复微阵列数据的不同方法和获取独立的基因表达重复测量的重要性。

图像分析、芯片数据的表示和标准化

三个互相关联的部分——图像分析、标准化和芯片数据的表示与评价——论述了从单个的微阵列实验中采集和处理数据。虽然主要关系到从使用双色玻片微阵列杂交实验中得到的数据，但是许多概念同样与来自 Affymetrix 阵列和尼龙滤膜大型阵列的数据有关联。

PART Ⅱ　图像分析

有多种方法用于区分噪声与信号和进行背景校正，该因素对于数据输出有重要的影响。文中介绍了可以适当地扣除背景读数，以及尝试将微阵列上的“点”和垃圾加以区分的自动化程序。

PART Ⅲ　标准化

PART Ⅲ中介绍了检查和消除微阵列数据中系统性偏差的方法。使用的 RNA 量的差异和荧光染料掺入效率的系统性改变常常会导致从竞争性杂交的微阵列中获得的数据产生偏差（Tseng et al. 2001；Yang et al. 2002）。

PART Ⅳ　芯片数据的表示和评价

这部分讨论使用图形化的显示来查看和评价微阵列实验的结果，并且说明了如何描述来自单张玻片阵列的数据，以使得对数据质量的快速评价更加容易进行。

511　微阵列数据库

在一次微阵列实验中，以 10 μm 的分辨率扫描 20 000 个探针将产生 3 千万个离散的数据点，如果以 Tiff 文件的形式将之存储，要占用大约 60Mb 的磁盘空间。因此，即使是偶尔使用微阵列的用户，也必须寻找到有效的存储和分析数据的方法。

PART Ⅴ　LIMS、数据库和数据管理

这部分讨论了微阵列数据库的规范，并介绍了三个著名的公共数据库：AMAD、SMD 和 GeneX。AMAD 和 SMD 分别用于小规模和大规模的项目任务，这两个数据库的特点强调了选择数据库时要对预期的需要和容量进行筹划。

对来自多张玻片阵列的数据进行分析和可视化

微阵列实验能够揭示协同基因表达的模式（patterns of coordinate gene expression），从而反映出意料之外的生物学复杂程度。在血清刺激的成纤维细胞中的伤口愈合和组织重塑（Iyer et al. 1999），以及癌症分子亚型的确定（Golub et al. 1999；Aliza-

deh et al. 2000; Dhanasekaran et al. 2001; Hedenfalk et al. 2001; Kihara et al. 2001）都是很好的实例。在这两个例子之中，对协同基因表达的复杂模式的识别依赖于一种对数据集中的高级结构进行识别的方法。

PART Ⅵ 聚类分析与显示

Cluster 和 TreeView 这两个程序是微阵列分析的支柱（Eisen et al. 1998）。这部分内容包含了帮助研究人员获取和使用这些程序的信息。

PART Ⅶ 多维尺度法和自组织图

其他两种非监督的方法——多维尺度法（multidimensional scaling，MDS）和自组织图（self-organizing maps，SOM）——被用于微阵列数据的整理排序和可视化。这部分内容帮助研究人员获取和使用这些程序。现在这些程序可以在 http://arrayanalysis.nih.gov 上获取。

只要存在引导初始分析的数据，计算方法就可以"训练"一种算法来识别数据中的模式，这就是所谓的有监督学习程序（Brown et al，2000; Khan et al 2001; Pomeroy et al. 2002），它能够比非监督的方法更加有效地分析特定的数据集。但是，在本书写作时，将这些方法应用于从微阵列实验中获取的复杂数据集才刚刚开始，这里将不涉及这些内容。

基因注释和微阵列信息公共存储的方法

微阵列数据挖掘的一个关键部分是有效地链接到包含感兴趣的单个基因的信息的数据库的能力。在基因注释信息栏中讨论了对基因列表进行注释，迅速将生物学信息与感兴趣的差异调节基因相联系时可供的选择。信息栏中还介绍了基因本体组织［gene ontology（GO）consortium］在使得基因注释更加容易进行方面所作出的努力。

在建立可供研究人员之间共享的基因表达数据库方面，微阵列界已经走出了第一步。关于研究人员之间共享微阵列数据的信息栏中，介绍了微阵列基因表达数据库（MGED）组织所作的努力，以提供一种结构化的方法来记录和共享微阵列数据。MIAME（minimal information about a microarray experiment，微阵列实验的最低限度信息）方案的目的就是提供一种可靠的方法来记录微阵列实验的"材料和方法"。

512

虽然用于存储和挖掘微阵列数据的生物信息学工具的发展非常迅速，但研究人员必须承认，当前还没有处理微阵列数据的"最佳解决方案"。在用于研究微阵列的制作和应用的设备中，一个常见的尴尬境地是低估该问题的规模。然而，一旦使用的设备能够用来生成微阵列，并且方案"正常工作"，拥有者就会急于使用它们。对几百个不同组织样品的表达谱进行比较与在少量样品中简单地识别一些差异表达的基因相比，需要大量更加复杂的工具。

一种解决方案是采取逐步增长的方法来应对数据存储和分析程序的实现。尽管这种方法的风险在于经常使得设备处在"紧随"数据问题的境地，但它是对付从一开始时就面临着令人畏缩的挑战的有效方法。非常重要的一点是，建立一个计划，采用第一个版本的解决方案来处理最为迫切的问题，同时认识到随着设备规模的扩大和经验的增长，

还将会有第二版和第三版的解决方案等。数据必须以持久的方式进行存储，并且在记录关键实验细节时要知道，对这些细节的系统性描述是在不久的将来发表研究结果的先决条件。

千万不要绝望！正如第一个微阵列硬件制造者所做的开创性努力已部分让位于更容易应用的商业化解决方案那样，全面的、用户友好的微阵列生物信息学软件的开发将会迅速出现。但是，需要记住的是，对基因表达复杂模式的分析还处在初级阶段，研究人员必须快速行动起来，要超越仅仅借助于鼠标的“指向和点击”的态度来面对数据分析。

513

PART Ⅰ 微阵列表达实验的设计

Yee Hwa Yang，Terry Speed（加州大学统计系，伯克利，加利福尼亚，94720）

导言

良好的实验设计可以简化分析，并且使得数据的解释更加有力。设计必须将这些目标和微阵列的花费与可利用性，以及可供实验和重复之用的 RNA 数量等约束因素进行平衡。在 cDNA 微阵列实验中，主要的设计选择是两个样品是直接在一张玻片阵列上进行比较，还是通过第三个样品如一个公用的参考进行间接比较。

基于微阵列的表达分析本质上是比较性的。样品通过与点样在同一张玻片或者不同玻片上的探针进行杂交来进行比较。因为这类比较实验的设计原理并不新颖（Kerr and Churchill 2001a），比较实验的设计和分析可以在经典的统计学框架内进行。

需要用实验来处理单因素问题时，实验方法的选择可能是简单明了的，然而研究人员常常希望在一次实验中考察一个以上的问题。因为每个问题对实验设计都会有影响，所以要设定问题的优先顺序，并决定要得到各种不同的结果所采用的最佳方法。可以根据对问题的理论分析，以及比较先前的实验中使用不同设计时获得的数据来深入了解实验设计方法对于微阵列分析的分辨能力的影响。

设计的选择受到实验优先级的影响。例如，假定从患者获取了大量的肿瘤样品（P），同时得到了对应的正常组织（N）。如果主要目的是寻找在肿瘤中发生上调或者下调的基因，那么最佳的方法是将从肿瘤细胞和正常细胞中抽提的 RNA 在同一张玻片上进行直接比较。相比之下，如果目标是要区分肿瘤的亚型（Alizadeh et al. 2001），那么将一系列肿瘤与一个公用的参考 RNA 进行比较的实验可能会更好一些。

在进行 cDNA 微阵列实验的设计时，为了将有时模糊的生物学问题转换成为合适的统计学框架，必须解决某些普遍性的问题（请见下面的信息栏）。正如上面所讨论的，确定实验的主要目标是非常重要的。要对先前实验的结果进行分析，因为它们可能会表明发生差异表达基因的大致数量，从而会影响到获得统计学显著性所需要的玻片数量。用于验证微阵列实验的方法［如 Northern 或者 Western 印迹分析、反转录酶-聚合酶链反应（RT-PCR）、原位杂交］也非常重要，因为这会影响到分析中使用的统计方法，

以及决定样本的大小。例如，使用相对低通量的技术来进行验证，那么就需要微阵列实验结果有高度的确定性。反之，高通量的验证技术则更容易处理微阵列实验中的噪声。RNA 的来源（如组织样品或细胞系）会影响到可利用的 RNA 的量，进而会影响到实验中可能的重复玻片的数量。样品分离、mRNA 抽提和探针标记的方法也都会影响到重复的性质和数量（请见下文关于普遍性设计问题的讨论）。对照样点可以简单地用来检查杂交是否成功，或者在分析如数据标准化时是非常有用的。

514

设计有效微阵列实验时普遍性的考虑

- 实验的目的是什么？这些微阵列实验在整个研究计划中起着什么作用？这些实验能否有效地解决提出的问题？
- 提出的特定问题是什么，依此而言，最重要的比较又是什么？根据可利用样品的类型和数量，指定所提出的不同科学问题的优先顺序。
- 计划采用何种对照？必须要进行标准化吗？或者只是希望进行？
- 采用何种方法对微阵列实验的结果进行验证？是否有足够的实验材料进行 Northern 杂交？RT-PCR 是否优于其他方法，如免疫组织化学或原位杂交？
- 哪些资源可供利用，数量有多少？这里要考虑的因素包括样品的类型和可供使用的玻片的数量、需要的 mRNA 的量，包括诸如重复和染料交换等类型的对照。
- 额外的信息有杂交前的实验过程（即样品分离、mRNA 的抽提、标记和杂交）是否顺利进行。上游的实验过程能否及时产生微阵列分析所需要的全部试剂和生物材料。
- 下游分析。哪些类型的数据分析是合适/希望的？哪些类型是可供利用的？

设计选择

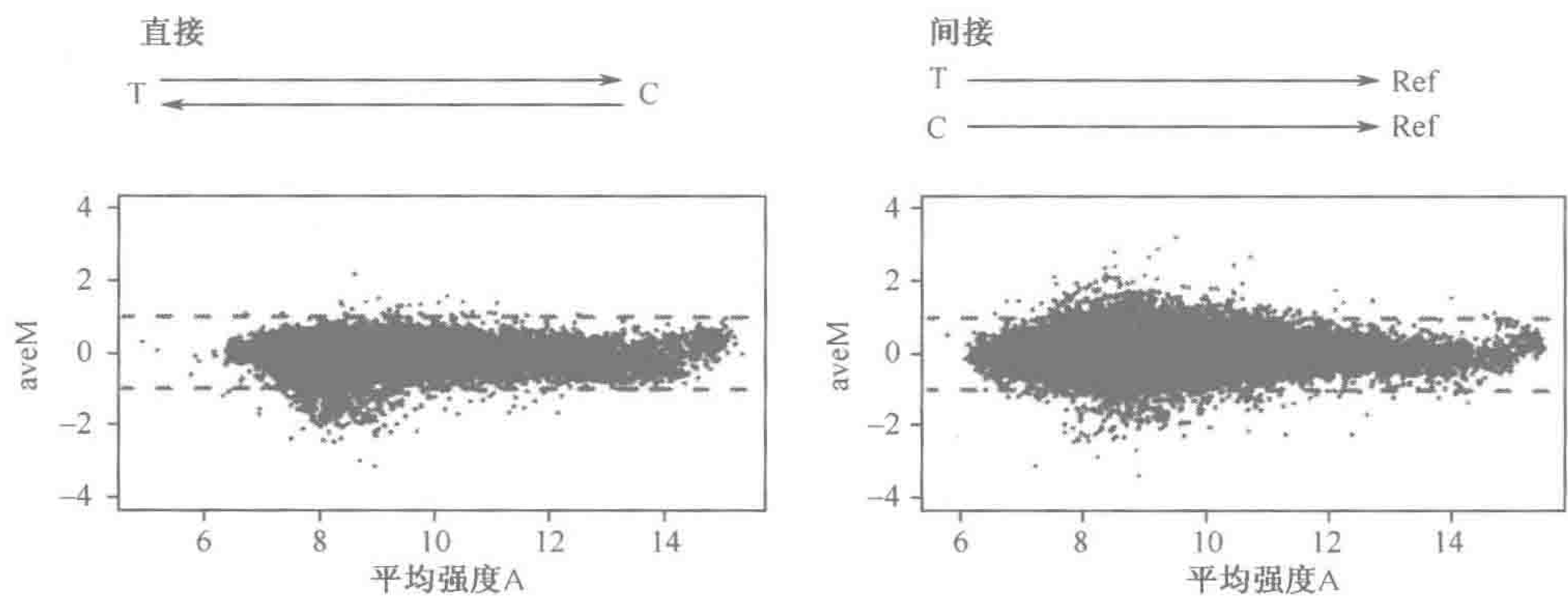
一旦确定了实验的性质和可利用的材料，在一些情况下会有一种设计方法优于所有的其他方法。例如，如果要将不同药物处理的细胞与未处理的细胞进行比较，那么很明显一个合适的设计是将没有处理的细胞作为实际的参考，并且全部的杂交都包含了处理的细胞集合与未处理细胞的比较。但是，对于大多数实验，可以设计出很多种看上去适用的设计方法，因此需要遵循一些原则以从这些可能的方法中选择出一种设计方法。

cDNA 微阵列设计的核心问题是在使用直接比较而非间接比较之间做出决定，也就是说，在各个玻片内部而非玻片之间进行比较。最简单的例子包含了处理 T 和对照 C。为了方便下面的讨论，“实验”和“对照”这两个术语泛指进行研究的细胞或者组织，与正常的或者未经处理的细胞或者组织之间的比较，例如，接受药物处理与未经处理的、野生型与突变体的（包括基因敲除或者转基因的），或者两种不同的组织（肿瘤与正常的）。

假定计划进行两次杂交，同时 RNA 的量不受到限制。在进行直接比较时，T 用 Cy5 标记，并且 C 用 Cy3 标记，两个样品（T-C）同时在两张玻片上杂交。对于任意一个特定的基因，可以得到关于该基因的对数比值 log(T/C)的两个独立估计。假如每次这种测量的方差是 $\sigma 2$，那么两次独立测量方差的均值为 $\sigma 2/2$（注意方差代表了单个测

量值，或者数据点围绕均值的分散程度，因此表示了数据的变异性）。

另一方面，如果采用了一个公用的参考，那么两次杂交将是 T-Ref 和 C-Ref。在这种情况下，任何基因的对数比值 log(T/C)可以通过公式 log(T/C)=log(T/Ref)－log(C/Ref)从两个独立对数比值的差当中得到。与先前一样，如果单个对数比值的方差是 σ^2，那么两个独立对数比值差的方差为 $2\sigma^2$。如下文所述，如果参考 RNA 来自于同一次抽提或者标记反应，那么两个分母就是相关的，因此方差可能会小于 $2\sigma^2$。总之，对于两次杂交，进行两次直接比较时每个基因的对数比值测量的方差为 $\sigma^2/2$，而间接比较时的方差为 $2\sigma^2$，后者比前者要大 4 倍。直接比较的较高精度是直接和间接测量的主要差别（图 7-1）。实际上，很少会出现 4 倍的差异，但为了简化起见，在计算时我们假定所有的样品是独立的。由于靶 mRNA 可能更多的是技术重复而非生物学重复，所以这种独立性假设往往不能满足（见下面关于重复的讨论）。



515

图 7-1 对应用直接和间接方法得到的处理与对照样品中相对基因表达水平估计的比较

相对表达的测量（aveM=平均的 $\log_2$ 比值）对两个样品中每个基因的平均强度（A）作图。尽管在两种设计中都使用了两张玻片，从直接比较（左图）获得的平均对数比值明显要比那些从间接比较（右图）中获得的结果分散程度窄一些。红色的点代表了一个已知差异表达的基因。很明显它湮没在公用参考作图的噪声中（右），但是在直接比较中它突出得更加明显（见彩图）

要使得提出的设计能够被接受，在玻片的数量和可供利用的 mRNA 的量给定时，实验在实际上必须是能够实现的。需要明确的是，推动实验的科学问题要能够被所选择的设计所解答，并且重要的问题比重要性低一些的问题回答得更为准确。下面的一些例子表明直接和间接比较相结合常常是对一个设计问题较实用的解决方法。

设计原理示例

例 1：多种来源的 mRNA 之间的比较

考虑三种不同来源的 mRNA，并且假定全部的成对比较具有同样的重要性。这种类型实验的一个例子来自于对大脑不同区域（A、B、C）之间表达差异的研究。表 7-1 中列举了一些可供选择的实验设计，这里“Ref”表示一个 mRNA 公用参考来源。

表 7-1　对多种来源的 mRNA 比较时可供选择的设计

	Ⅰ公用参考	Ⅱ直接比较	Ⅲ公用参考
	A B C Ref	A C B	A B C Ref
玻片的数量	$N=3$	$N=3$	$N=6$
均方差	2v	2v/3	
需要的单位材料数量	A=B=C=1	A=B=C=2	A=B=C=2
均方差		2v/3	v

对均方差进行了两次计算：第一次（浅灰色）用于比较使用相同数量玻片的设计Ⅰ和设计Ⅱ；第二次（深灰色）用于比较使用相同数量材料的设计Ⅱ和设计Ⅲ。

注意到因为所有的成对比较都具有同等的重要性，所以主要的科学约束为它们是可以估计的。根据不同的物理限制，如玻片的数目或者可供利用的样本数量，会作出不同的设计选择。尽管设计Ⅰ和设计Ⅱ中使用了同样数量的玻片，但是在设计Ⅱ中每种样品均使用了双倍的量。设计Ⅱ和设计Ⅲ中使用了相同数量的样品，即使是实验重复进行，设计Ⅲ的精确度却比较低。从直接比较所有成对组合的设计Ⅱ中能够得到对三个区域之间转录差异最精确的估计，并且一个额外的好处是节省了所需要玻片的数量。如果从

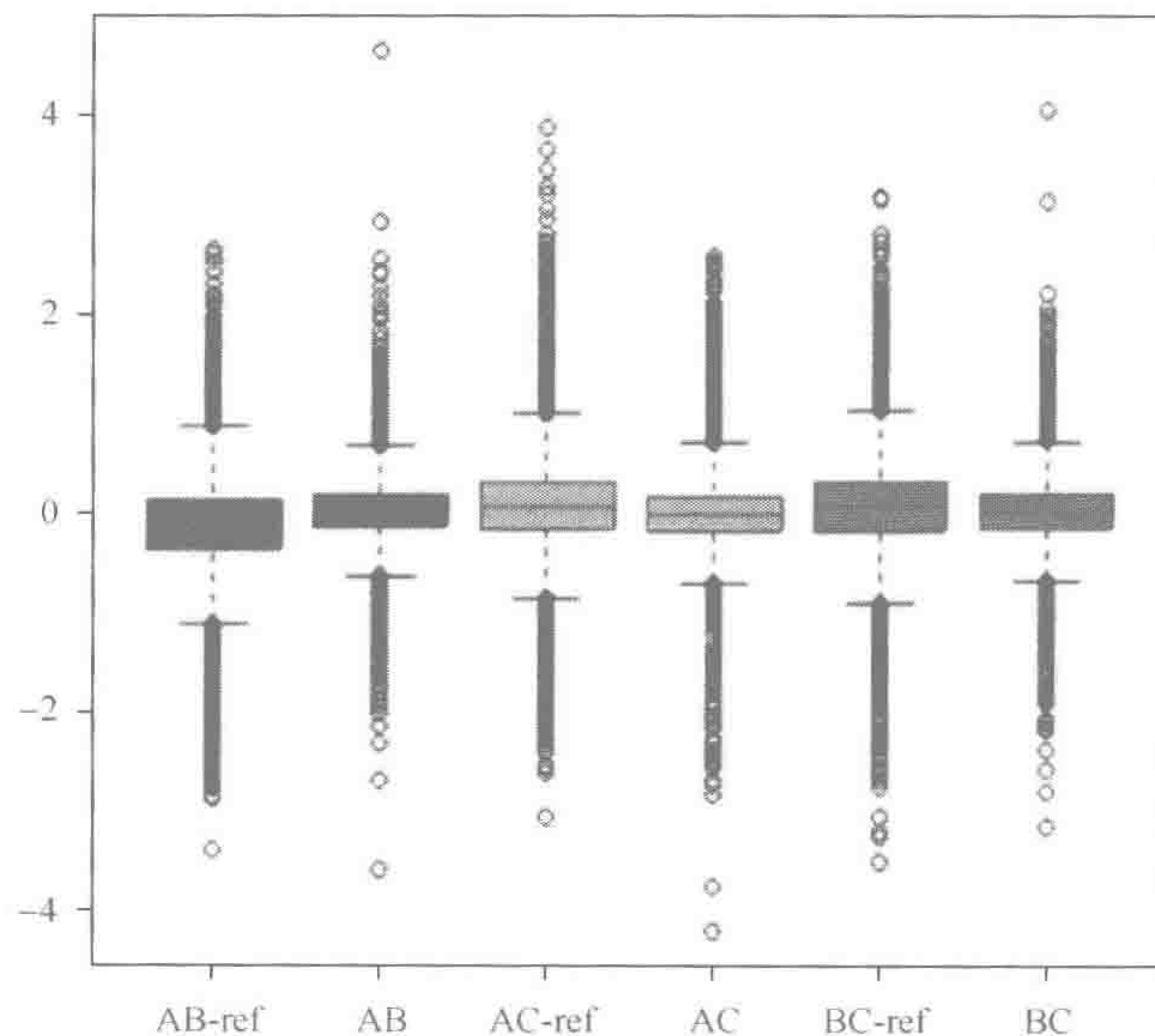


图 7-2　箱线图示例样品间的直接比较要比间接比较具有较高的精度

每对样品表达比值的 $\log_2$ 作图，带有颜色的矩形盒表示四分位距（IQR），定义为 75%（上四分位数）和 25%（下四分位数）之间的差异。对箱线作图方法的介绍，见第 4 章

A、B 和 C 三个区域中分别只能获取一个样品，那么设计Ⅰ是唯一可能的选择。图 7-2 中显示了这些方法（直接和间接的）中每一种的相对精度。

517

当对超过三种来源的 mRNA 进行比较时，实验设计成为更加复杂的问题。参考设计（表 7-1，设计Ⅰ）适用于大量样品的横向比较。当超过四个进行比较时，从实验花费和材料需要数量的角度来看，全部配对（all-pairs）设计（设计Ⅱ）变得越来越具有挑战性。在对四种来源的 mRNA 进行比较时，需要 6 次配对比较；五种来源需要 10 次，依此类推。当实验的后勤供应越来越困难时，就必须考虑一些其他类型的、需求玻片数目比较少的实验设计，包括环状（loop）设计（见下面的讨论）（Kerr and Churchill 2001）。

例 2：时间序列实验

在时间序列实验中，最佳设计方法的选择关键取决于实验中时间节点的数目。表 7-2 中列举了一系列微阵列时间序列实验中可供选择的设计。设计Ⅱ涉及了进行连续时间节点之间的杂交，而设计Ⅰ中以 T1 作为公用的参考。当实验的主要目的是研究 T2、T3、T4 和最初的时间节点 T1 之间的相对变化时，设计Ⅰ是比较好的选择。然而，如果使用设计Ⅱ，从一个到另一个时间节点之间更加细微的变化会更加容易地被检测到。

表 7-2　时间序列实验中六种可能的实验设计

	时间序列实验中的设计选择	t vs. $t+1$			t vs. $t+2$		t vs. $t+3$	Ave
		$t1t2$	$t2t3$	$t3t4$	$t1t3$	$t2t4$	$t1t4$	
$N=3$	（Ⅰ）T1 作为公用参考 T1 T2 T3 T4	1	2	2	1	2	1	1.5
	（Ⅱ）直接比较 T1 T2 T3 T4	1	1	1	2	2	3	1.67
$N=4$	（Ⅲ）公用参考 T1 T2 T3 T4 Ref	2	2	2	2	2	2	2
	（Ⅳ）T1 作为公用参考＋更多的比较 T1 T2 T3 T4	.67	.67	1.67	.67	1.67	1	1.06
	（Ⅴ）环状设计 T1 T2 T3 T4	.75	.75	.75	1	1	.75	.83
	（Ⅵ）T1、T2 作为公用参考 T1 T2 T3 T4	1	.75	1	.75	.75	.75	.83

表中给出了全部配对比较相应的估计方差和每种设计的平均方差。

如果设计不受可以进行的杂交数量的限制，对不同实验设计之间的选择变得更加困难。设计Ⅲ给出了一种使用非时间节点作为公用参考的公用参考设计方法。设计Ⅳ是一个类似的公用参考设计，使用 T1 作为公用参考，同时附带了 T2、T3 以及 T3、T4 之间的直接杂交。设计Ⅴ和Ⅵ每种均提供了不同的直接和间接比较的混用，导致一些比较的准确性较高，而其他比较的准确性却较低。对于设计Ⅴ和设计Ⅵ的选择明显取决于样品比较的重要性，因为这两种设计中比较的平均方差是相同的。当连续时间之间的比较要比两个彼此分开的时间节点更加重要时，设计Ⅴ因是一种环状设计，要比设计Ⅵ好。如果这种重要性的优先级正好相反，设计Ⅵ则更好。 **518**

例 3：2×2 因素实验

先前的例子中涉及了单因素或者单向设计，其中单个因素（例 1 中的组织、例 2 中的时间）具有三个或者更多的水平。同时，考虑两个或者更多的因素，并且每个因素具有两个或者更多的水平时，出现了一类更加复杂的设计类型。这些称为因素实验，用来研究单个因素所导致的，以及那些由两个或者更多因素的共同作用所导致的表达差异。任何关于因素实验的讨论必须包含对因素之间相互作用认识的研究。因素实验由 Fisher 在 1926 年提出，并与他的合作者 Yates（1937）进行了广泛研究。关于更为现代的处理方法，见 Box 等（1978）。下面介绍一个 2×2 因素的例子。

例 3a：细胞激素的组合（图 7-3）

假定对细胞系有两种不同的处理方法，例如，向生长培养基中加入不同的生长因子。如果用 C 表示从未处理（对照）细胞中得到的 mRNA，A 和 B 分别表示从用第一种生长因子和第二种生长因子处理的细胞中得到的 mRNA，那么 AB 表示同时使用两种生长因子处理过的细胞（图 7-3 中描述了交互关系）。处理 A 对于基因表达的影响可以通过研究在第二种处理不存在的情况下，某个特定基因在来源 A 和 C 中的相对表达来进行估计，这里处理 A 的影响可以用 log(A/C)来估计。另外，在 B 存在的情况下，因素 A 的影响使用 log(AB/B)来估计。log(AB/B)－log(A/C)＝log(AB×C/A×B) 的差值称为对数尺度上两种处理的相互作用。我们可以把相互作用理解为 A 引起的基因表达差异对 B 是否存在的依赖程度的测量。使用的测量方法在 A 和 B 中是对称的。

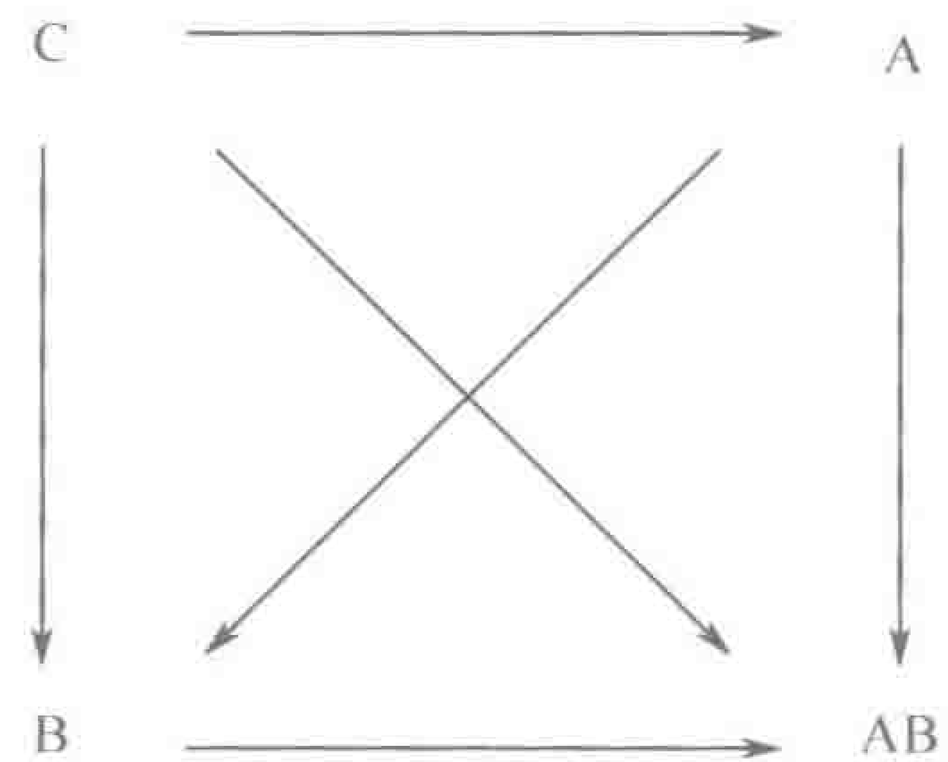


图 7-3　在 2×2 因素实验中可能的比较

例 3b：细胞类型和处理之间的交互作用 **519**

假设有两种类型的细胞，一种特定类型 T（如肿瘤细胞）和另一种对照类型 C（如来自同一器官中的正常细胞），两种细胞类型均接受辐射处理（RT 和 RC）。在这种情况下，感兴趣的是确定受辐照的肿瘤细胞 log(RT/T)与受辐照的正常细胞 log(RC/C)相比较，哪些基因的差异表达情况不同，表示为 log(RT/T)－log(RC/C)。

例 3c：组织-发育阶段之间的交互作用

对两种类型组织 L 和 M 之间的基因表达，例如，在两个不同的时间 Te（胚胎期）和 Tp（出生后）进行比较。在这个实验中有 4 种可能的 mRNA 样品（见图 7-3）：C 为处在胚胎期 Te 的区域 L；A 为处在出生后 Tp 的区域 L；B 为处在胚胎期 Te 的区域 M；AB 为处在出生后 Tp 的区域 M。可能会期望在 A-C 比较中差异表达的基因应该和 AB-B 比较中的基因大体上相似，因此对于某个特定基因，注意力应该集中在差值 $\log(AB/B)-\log(A/C)$上。

这些设想的差异将会如何影响到实验的设计呢？假设有四种来源的 mRNA，并且主要关心的是发现对数比值 $\log(AB/B)$和 $\log(A/C)$的差值较大的基因，也就是说，主要的目的是发现相互作用较大的基因。可以通过多种方法来实现这个目标，表 7-3 中给出了一些设计（Glonek and Solomon 2002）。要注意这里所有的设计都包含了 6 张玻片。在设计Ⅱ和设计Ⅳ中，相互作用项的方差最小，但是在设计Ⅳ中，对 A 的主要作用即使是估计也是不行的。设计Ⅰ对于相互作用的估计是最差的，但是使用了较少量的 mRNA：每种来源只需要两个单位，与此形成对比的是全部其他设计中都需要三个单位。假如任何物理约束都能够得到满足，所喜好的设计将取决于考虑到交互作用时主要作用的重要性程度。通常，设计Ⅱ（或者是和它互补的方法，也就是水平而非垂直的染料交换）可能是首选的设计，它为所有的比较都提供了好的准确性，尽管它对一种主要作用的准确性要比其他的高。

表 7-3　2×2 因素试验的四种可能设计

设计选择	间接方法	直接和间接方法的平衡		
	(Ⅰ) A　B　A,B C	(Ⅱ) C　A B　A,B	(Ⅲ) C　A B　A,B	(Ⅳ) C　A B　A,B
玻片的数量	$N=6$			
主要作用 A	0.5	0.67	0.5	n. a.
主要作用 B	0.5	0.43	0.5	0.3
交互作用 A,B	1.5	0.67	1	0.67

仿 Glonek 和 Solomon（2002）。

520

一般设计问题

重复实验

为什么要对玻片进行重复实验呢？这里一个简短的解释是：①通过累加统计减小变异性；②可以使用习惯的统计学方法。其实，重复实验能够进行平均，而平均值一般会

比它们的组成项具有较低的变异性。在单个基因的水平上，cDNA 微阵列系统的噪声相当大。在一次杂交中一个基因的差异表达可能有 4 倍，而在第二次独立的杂交中可能只有 1.3 倍的变化，并且在第三次中为 2 倍。为了确定两个 mRNA 样品之间哪些基因发生了差异表达，并且对该结果不是假阳性具有一定的置信度，重复实验是必要的（Lee et al. 2000）。缺少重复实验严重地限制了利用统计推断方法来确定一个给定的强度对数比值是否与 0 有显著的差异。特别的，重复实验对于估计不同玻片之间对数比值的变化是必需的。试图仅使用来自一张玻片的数据去估计对数比值的显著性将依赖于未经验证的模型假设，并且通常无法考虑最重要的差异来源，即玻片间的差异。如果使用重复的玻片，就可以应用定义相称的、标准的统计学方法（Dudoit et al. 2001）。重复实验通常包含了统计学上随机的 mRNA 制备物样品，并且在这些条件下，从样品外推到同样的总体是正确的。在这个意义上，重复实验与将结论加以推广是密切联系的。

图 7-4 中显示了在所有的重复玻片中对数比值 $\log_2(\mathrm{KO/WT})$ 的均值，对同样平均处理的总体强度 $\log_2(\sqrt{\mathrm{KO \cdot WT}})$ 的作图。每张重复玻片包含了来自一只不同实验动物的 mRNA，都与同样的参考 mRNA 进行同时杂交（Callow et al. 2000）。绿色的点对应于在两种来源的 mRNA 中已知差异表达的 8 个基因。随着样本量（本例中为老鼠的数量）的增加，成群的数据点在水平轴上的分布发生收缩。这使得区分真实的改变和趋于 0 的随机变化变得容易。注意到仅有一个重复时，数据点群在 $\log_2$ 尺度上的分布

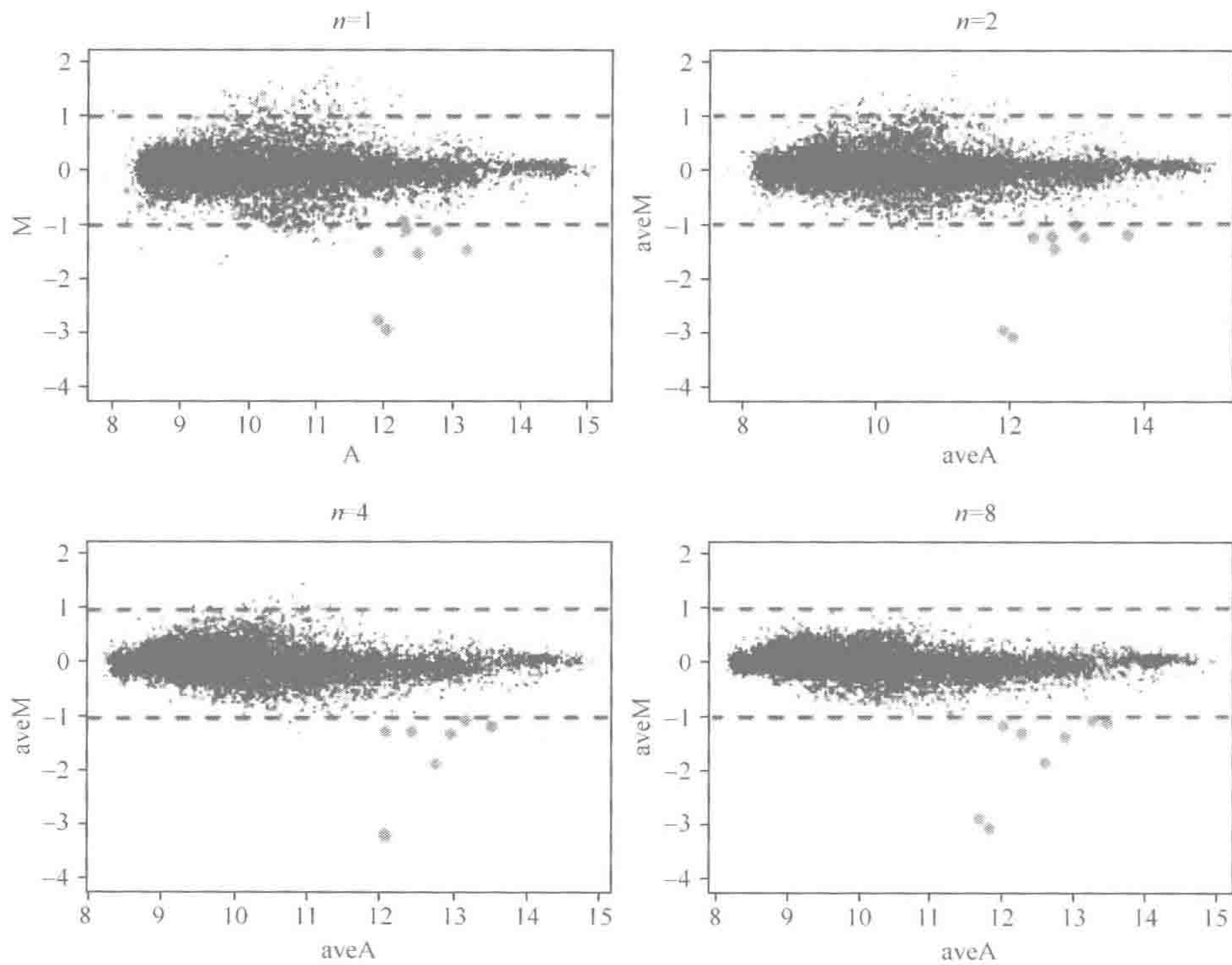


图 7-4　所有重复玻片的对数比值 $M=\log_2(\mathrm{KO/WT})$ 的均值，对同样平均的总体强度 $A=\log_2(\sqrt{\mathrm{KO \cdot WT}})$ 的作图

521

绿色的点对应于在两种来源的 mRNA 中已知差异表达的 8 个基因。图中所示的重复玻片数量为 $n=1$，2，4 和 8

超过±1，或者在任一方向上为两倍。与此相比，当 $n=8$ 个重复时，数据点群在 $\log_2$ 尺度上主要分布在±0.7 的范围内，在任一方向上约为 1.6 倍。

如果 8 组重复实验数据真正是独立的，这里并不为真，因为全部 8 只老鼠共用了同一个参考 RNA，那么可以预期数据点群的分布还会有更大的压缩：在对数尺度达到±0.35,或者在任一方向上为 1.3 倍的变化。

不同形式的重复实验

如同上面所解释的那样，并且与统计学的惯例相一致（Fisher 1926），重复实验是比较性微阵列实验一个非常必要的特点。重复实验有许多种不同的方法，差别在于数据可以看成是独立的程度以及实验样品对总体的代表程度。假使同一个研究人员，使用相同的设备和方案，几乎不变地进行重复杂交，并且常常在相同的时间，那么重复数据必将不可避免地具有许多共性。下面列举的大多数重复方法关注的是靶 mRNA 样品。

- **重复样点**。许多研究人员，常常在邻近的位置上在每张玻片上对 cDNA 进行重复点样。这种方法提供了有价值的质量信息，因为重复点的强度或者相对强度一致性的程度是一个很好地反映质量的指标。但是，由于重复样点位于同一张玻片上，特别是邻近的样点，实验条件即使不完全相同也会大体相似，因此这些成对数据不能看做是相互独立的。通常，重复样点总体上的一致程度要明显高于在不同的重复玻片上观察到的相同的点。尽管对重复样点的对数比值进行平均值是适当的，但是物理位置上的毗邻导致得到的信息比从真正独立的成对样点中得到的信息的用处要小。因此，如果要在同一张玻片上点样重复样点，那么最好将它们分开而不是相邻，因为这样可以比较好地看出芯片上不同部分之间的变化。
- **重复玻片：技术上的重复**。技术重复这个名词用来表示使用相同的靶 RNA 制作的玻片。图 7-5 中示例了与使用独立的 RNA 制备物相比（M1 vs. M3），当使用相同的 RNA 制备物时（M1 vs. M2）带来的偏差。这类结果证明了技术重复不能够提供为充分利用平均处理带来的优点所必需的独立性：在技术重复样本中，共有的系统特性（偏差）在平均后仍然存在。

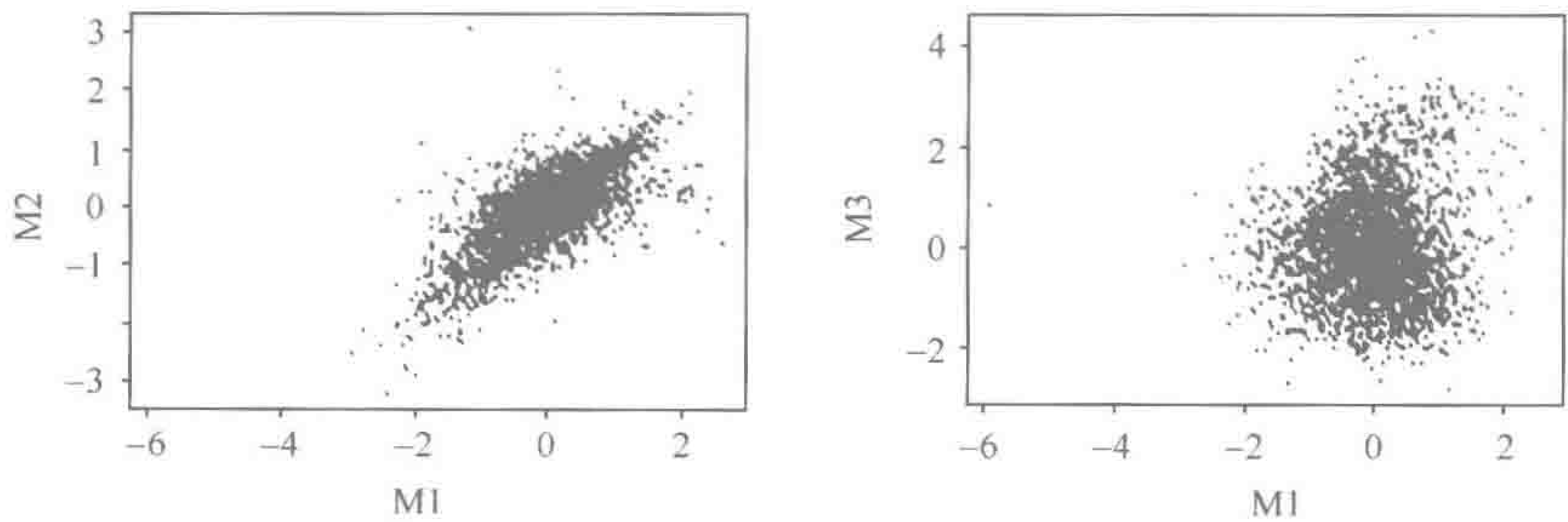


图 7-5 使用相同（左图）和不同的（右图）mRNA 制备物制备重复玻片的对数比值的散点图

全部三个实验（M1、M2、M3）包含了自我杂交。注意到看上去差异表达的基因在 M1 和 M2 中是一致的（相关的），因为它们倾向于分布在对角线上。与此形成对比的是，在 M1 和 M3 的对比中，差异表达的基因是随机分布的，说明了单独制备的 RNA 样品的独立性

- **重复玻片：生物学重复，相同的来源**。生物学重复这个名词指应用从相同的组织或者细胞系的不同样品独立制备的RNA进行杂交。在许多情况下，这是最为方便的重复形式。假定对于不同抽提的RNA分别进行标记，从这种方法得到的实验结果将是这种重复形式之下最接近独立的。但是，如果使用了一个公用的参考，那么分子来自独立的生物学重复的对数比值仍然会显示出一定的相关性，因为它们在分母上具有共同的、来自同一批次的参考mRNA。
- **重复玻片：生物学重复，不同的来源**。生物学重复这个名词也可以指应用不同物种或者不同版本的细胞系制备的靶RNA来重复玻片。这种形式的生物学重复与上面讨论的生物学重复在本质上是明显不同的，而且在测量中会有更大程度的变异。例如，利用一群老鼠进行的实验中，不同老鼠的激素和免疫系统必然会处于不同的（波动的）状态，或者会有显示不同程度的炎症反应的组织等。这种重复形式本身所固有的较大的差异使得重复看上去是不必要的，而且会使得真实的表达差异难以区分。但是，从将结论推广的角度来看，这对于一些实验而言是合适形式的重复。

一个给定的实验中所使用的重复类型将会影响到准确性，并且对得到的实验结果的推广会产生影响。在大多数情况下，一个研究人员会想利用生物学重复来验证结论能否推广，而利用技术重复来减少这些结论中的变异。假定经常有多种可能形式的技术上和生物学上的重复，研究人员需要在实验和花销限制的前提下，根据对需要多少数量某种特定类型重复这个问题的回答来作出决定。例如，希望作出的结论适用于某个特定培育系的所有老鼠，必须进行包含了多只老鼠的实验，最好是这类老鼠的一个随机样本。如果根据从单只老鼠中得到的结果外推至该培育系的全部老鼠，即使是使用了多个mRNA抽提物，也有着广为人知的、与此相关联的风险（Freedman et al. 1998）。

合并的mRNA与未合并的mRNA相比较

523

在许多情况下，为了有足够量的样品进行一次杂交，往往需要将许多相似来源（如老鼠胚胎）的mRNA合并在一起。但是，假定mRNA的合并不是必须的，那么仍然需要合并吗？假定一个实验中要使用4只处理的老鼠和4只对照用老鼠，并且每只老鼠都能够提供足够量的mRNA进行单次杂交。研究者是否要将来自4只处理老鼠的RNA合并到一起，并且对作为对照的老鼠进行同样的处理，然后使用合并后的mRNA的子样进行4次实验？或者作为替代，研究人员是否要进行4次独立的处理-对照比较，然后将得到的对数比值结果进行平均？任何一种方法都没有多大问题，但是在缺乏数据的情况下，对在没必要进行合并时而进行合并的做法似乎没有很大的争议。不进行合并能让较大的差异得到显示，并因此增强平均处理的作用。

功效和样本量的确定：差异性表达

已知重复的重要性，并且选定了适合于当前研究实验中需要使用的重复形式，那么一个重要的实际问题是决定样本的大小，或者说需要使用玻片的数量。通常，研究人员需要阐明下述问题以进行功效计算：

- 单次测量的方差。

• 要检测的效应的大小。
• 可以接受的假阳性率。
• 检测到具有特定（或者大于）大小效应时所希望的功效或者概率。

研究人员可能会问：为了有90%的机会能够检测到基因表达中两倍的差异，需要进行多少次杂交？为微阵列实验确定最佳的样本量非常困难，因为不同的基因间相对表达水平差异在各次杂交中变化非常大。即使预先知道每个基因特定的变异，但这事实上是不可能的，研究人员通常也不能事先指定哪些基因的变化是重要的。发现这些基因常常正是微阵列实验的目的所在。因此，缺少用于标准的功效计算中两个重要的成分：

• 在基因水平上单个对数比值测量结果的方差 σ^2。
• 在同一水平上感兴趣的效应的大小。

能做些什么呢？假如我们知道差异表达基因的方差 σ^2 就能回答这个问题。克服这种僵局的一个方法是根据过去应用该实验系统的经验，赋给 σ^2 一个不是太小的值（确定该值始终是进行功效计算所必需的）。比方说，不同基因方差的中值或者上四分位值（upper quartile）。使用上四分位方差进行功效计算，研究人员就能够确认除了全部基因中25%的变异最大的基因以外，他们所使用的重复数量赋予了他们一定的功效，来检测大于某个规定值的差异性表达。

仍然存在一个难题使得我们无法举出一个简单的例子。研究人员通常不是孤立地考察一个特定基因可能的差异性表达，而是要在成千个没有发生差异性表达基因的背景中，检测可能来自于某类基因中的一个基因是否有表达差异。相对于我们要寻找证据的差异表达了的基因，其他基因形成了一个难以克服的背景变化水平，并且它们的存在使得功效计算更加困难。

对功效和样本量的讨论使得要在功效和假阳性率之间折中处理这一问题。尽管微阵列实验结果的验证具有常规的方法，是一个标准的统计问题，但是由于要检测大量的基因，因此会有特殊的、意料之外的变化。在用于微阵列实验的RNA样本量不足，但是验证方法既直接又相对花费较少的情况下，只要在验证过程中从假阳性里挑出真实的结果并不是很难，研究人员可能会愿意接受较高的假阳性率。在这些情况下，可以减少重复微阵列的数量。

观察全部亚型时样本量的确定

在确定样本量时要考虑的另一个问题是：为了检测到所有的分子亚型，需要对多少某个特定类型（如肿瘤）的样本进行分析？答案在部分上依赖于亚型的定义。在一种琐碎的意义上，每个样本能表示一种亚型，因此要检测的样本量是极大的。然而，与病理学家赞同疾病的组织学分类一样，尽管在个体样本之间存在着变异，描述性和推理性表达分析技术可以用来界定分子亚型。检测亚型时的另一个决定因素是随机抽样。例如，考虑一组卵巢癌的样本，其中的组织学亚型分布如下：浆液性囊腺癌（约70%的病例）、黏液性囊腺癌（约20%）、宫内膜样癌（约6%）、透明细胞癌（2.5%）和未分化细胞癌（1.5%）。假如对大量的病例进行重复随机抽样模拟，那么具有所有5种亚型的样本集合的数量将主要依赖于集合中包含的样本数量，并且在很大程度上由稀有亚型的分布频数决定：包含50个病例组成的样本集合＝36%的样本集合含所有的5种亚型，100

个病例时为72%，150个病例时为88%，200个病例时为95%，250个病例时为98%，300个病例时为99%。如果目标是实现比现有的亚型分类更加真实的分类，并且需要对这些集合进行进一步的分类，那么可能需要超过500个的样本。

染料交换的重复

在染料交换重复中，杂交的条件不变，只是在第二次杂交时将标记探针的两种染料互相交换，这有助于减少系统性的偏差（Tseng et al. 2001；Yang et al. 2001b）。大多数cDNA微阵列实验中，红色和绿色的强度会呈现出系统性差异，需要在归一化时进行校正。这种标准化处理并不能很好地将每块芯片上的每一个点的颜色差异都进行有效地校正，从而不残留颜色的偏差。只要不同玻片之间颜色的偏差是可重复的，对染色交换的成对玻片进行平均可以消除实验中产生这类偏差的倾向。相应的，只要有可能，建议进行成对的染色交换。作为选择，也可以使用随机的染料交换，这样可以包含随机性偏差。应该避免的是全部使用同一种染料，对需要直接比较的玻片进行重复。间接比较，如使用一个公用参考的比较，是基于芯片之间的差异的。由于这个原因，可重复的残留颜色偏差应该使用与染色交换时非常类似的方法来去除。

525

可扩展性和链接

研究人员常常希望在长时期内对本质上是任意的多种来源的mRNA进行比较。一个方法是全部实验中使用同一个公用参考的实验设计，这里的公用参考是从细胞系和组织中合并得到的"通用"参考RNA（见第3章，方案12）。尽管许多实验室制备他们自己的公用参考集合（见第6章，PART Ⅰ，方案1），一些厂商提供通用参考mRNA（http://www.stratagene.com/gc/universal_mouse_reference_rna.html）。公用参考为实验室内部和不同实验室之间的一系列实验提供了可扩展性。当一位研究人员被迫采用新的参考来源时，新的实验与先前基于不同的参考来源的实验进行比较就会变得困难。因此理想的公用参考是所有人都容易得到的，而且数量不受限制，并且对于很大范围内的基因都能够给出一个信号。在实际应用中，这些目标实现起来很困难。

当一种通用的参考RNA无法再得到时，必须进行额外的杂交，我们称之为链接实验（D. Lin et al. in prep.），用来关联无关的实验。更为普遍地，链接实验可以使得无关的实验被关联起来，额外杂交的次数依赖于所需要的结论的精度。假设在一个系列的实验中使用了参考R，然后在另一个系列中使用了参考R′。链接实验对R和R′进行比较，这使得对两种来源的样本可以进行比较，其中一个来源的样本与R共同杂交，另一个与R′共同杂交。然而，这种能力是需要代价才能实现的。将样本来源A与参考R共同杂交得到的对数比值与样本来源Z和参考R′共同杂交的对数比值进行比较，只有通过恒等式 $\log(A/Z)=\log(A/R)+\log(R/R')-\log(Z/R')$ 将A-R和Z-R′与R-R′联合起来。换而言之，交叉参考的比较方法包含了三个对数比值的联合，由于在这里 $\log(A/Z)$ 的方差是单个对数比值方差的三倍，因此会导致相应的准确性的降低。尽管如此，还是存在着值得进行交叉参考的时候，特别是当人们认识到链接项 $\log(R/R')$ 的方差可以简单地通过重复实验的方法减小到想要的程度时。这时恒等式中的链接项应该用全部重复中该项的均值来代替。

526

PART Ⅱ 图像分析

Yee Hwa Yang，Terry Speed（加州大学统计系，伯克利，加利福尼亚 94720）

导言

从被扫描的玻片上采集微阵列图像的方式对后续的分析有着很大的影响。在微阵列实验中，杂交后的微阵列用微阵列扫描仪扫描生成覆盖整个阵列的、反映红色和绿色荧光强度的大量像素集合（有关一些商业化微阵列阅读器/扫描仪的介绍，请见第 3 章，表 3-9）。荧光强度对应于两种靶分子与点样在玻片上的探针分子之间的杂交水平。这些荧光强度值以 16 位的图像形式存储，通常称为“原始”数据。

在过去的数年中，出现了许多用于玻片 cDNA 微阵列图像分析的软件包，包括商业化的和免费软件。这些软件包中有一些是从分析点样在尼龙膜上阵列产生的放射性信号的软件变化而来（关于对滤膜阵列成像的讨论，请见第 4 章）。其他软件是专门针对玻片微阵列而设计的。这些专门设计的软件包在其样点确定算法中利用了样点的规则排列特点以及两个荧光通道的信息，并且专门处理了荧光信号不同于由放射性产生的信号的特性。表 7-4 中列出了一些现有的微阵列图像处理软件。对扫描的微阵列图像进行处理可以分为三个主要的任务：

527

表 7-4 用于 cDNA 微阵列的图像分析软件包举例

名 称	单位/公司	网 址
ArrayVision	Imaging Research Inc.	http://imaging. brocku. ca/products/ARV. asp
Array-Pro Analyzer	MediaCybernetics	http://www. mediacy. com
AlphaArray	Alpha Innotech cooperation	http://www. alphainnotech. com/bio/lsmaster. html
ArrayWoRx	Applied Precision	http://www. api. com. /products/bio/arrayworxe. html
AutoGene 2. 5 ImaGene 4. 0	BioDiscovery	http://www. biodiscovery. com
DeArray	ScanAlytics	http://www. scanalytics. com/product/hts/microarray. html
GenePix 3. 0	Axon Instruments	http://www. axon. com
GLEAM	Center for Behavioral Neuroscience，Atlanta	http://www. cbn-atl. org/research/molcore. html
P-Scan/F-Scan	Analytical Biostatistics	http://abs. cit. nih. gov/
QuantArray	Packard Bioscience	http://www. packardbioscience. com/
ScanAlyze	Eisen Lab，LBNL	http://rana. lbl. gov/EisenSoftware. htm
Spot	CSIRO	http://www. cmis. csiro. au/iap/spot. htm
Visage HDG	Genomic Solutions	http://www. genomicsolutions. com
Virtek ChipReader	Virtek Vision International Inc.	http://www. virtekvision. com/ http://www. virtekvision. com/pdf/imageanalysis. pdf

- **定位或网格化**。是给阵列上的每个样点指定坐标的过程。这部分过程的自动化使得高通量分析成为可能。
- **分割**。将像素归类为前景，也就是位于点样的 DNA 样点内部，或者背景。
- **强度提取**。包括为微阵列上的每个样点计算成对红色和绿色的前景荧光强度（R，G）、背景强度以及可能的质量评价。

在微阵列图像分析中，玻片点样、靶分子杂交以及图像扫描的质量越高，图像分析程序就越容易在每个阶段准确地测量样点的强度，这一点非常重要。通常，每次点样过程中最后一张玻片的样点形态差异最大。

一个常用的微阵列成像系统中包含了激光共聚焦扫描装置，其中的激光用来激发荧光标记的靶分子。发射的光子被光电倍增管（PMT）放大，产生大量的电子，并由检测器收集为信号。在不同的扫描仪中，可以对许多设置进行调整，包括扫描速度、激光功率和 PMT 的电压。PMT 或者激光设置参数的增高可以使得信号和噪声都得到增强，而且增高的激光功率会通过光漂白作用损伤杂交的样品。

设置很高的 PMT 电压可能会使得像素饱和，也就是说，因为探测器寄存的信号为 $2^{16}-1$（16 位图像），可以记录的最大信号为 65 535。在实际应用中，将 PMT 的电压调节为使得最亮的像素恰好处在饱和水平之下。如果不这样设定，很强的信号会被低估。对在一系列的 PMT 设置下扫描后的图像进行比较，发现经过合适的归一化处理后，大多数基因的对数比值和排序保持不变（J. Rahnenfuhrer et al. 未发表结果）。

定位

微阵列图像的基本布局或结构是已知的，因为这是由微阵列点样仪进行点样时决定的。该信息用来帮助微阵列图像分析软件确定图像中的样点。图 7-6 中表示了一个典型的微阵列布局。为了在微阵列图像中确定样点的位置，也就是说，将一个理想的微阵列模型与扫描后的图像数据进行匹配，必须对许多参数进行估计，包括：

- 由于点样针位置的轻微变化而导致的网格或样点相对期望位置的位移（平移）。
- 个别样点细微的平移。
- 网格之间的行距和列距。
- 每个网格内部样点之间的行距和列距。
- 扫描后的图像中，微阵列在整体上的位置。

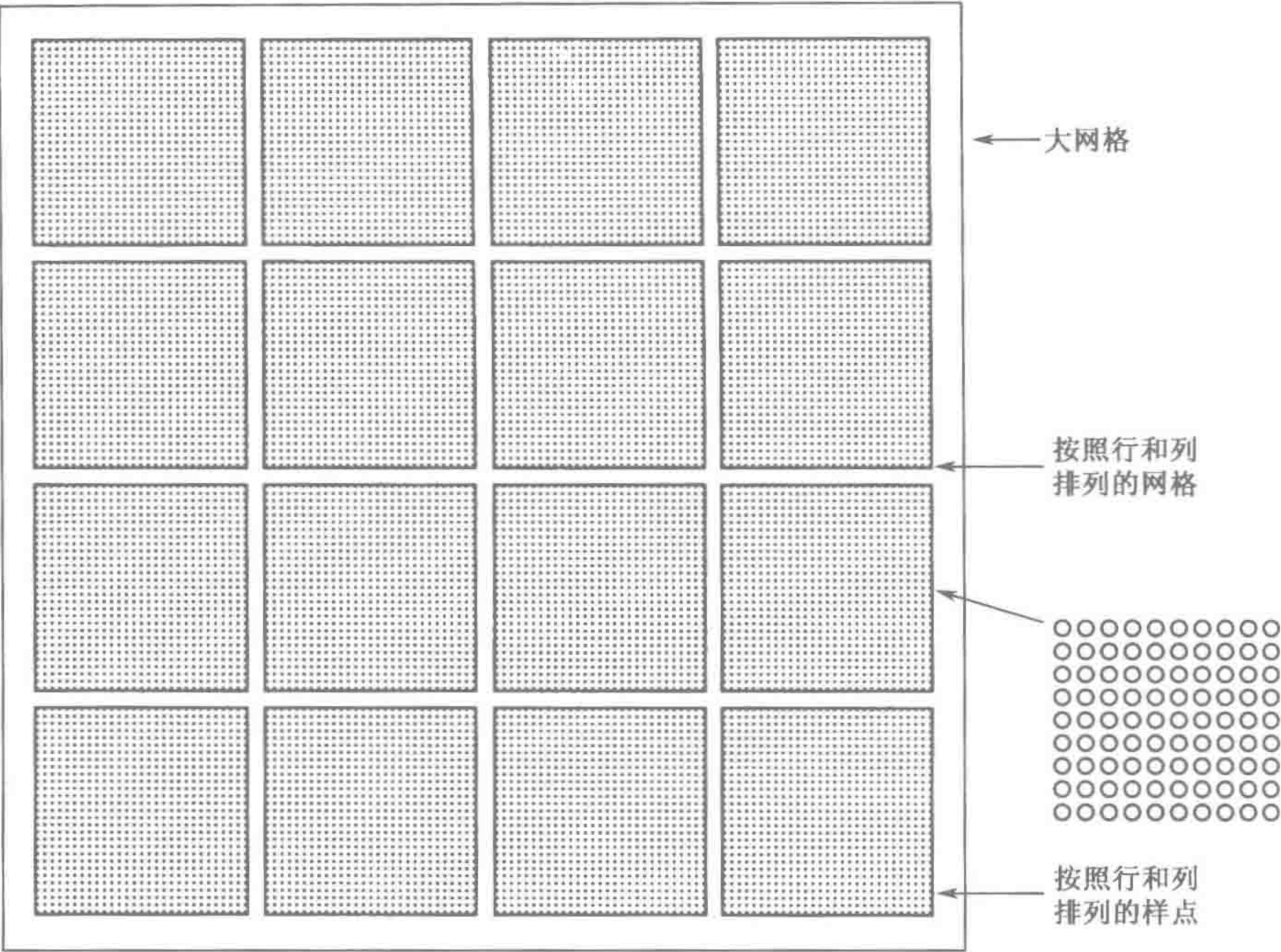
其他需要考虑的参数有红、绿通道图像的套准错误，图像中阵列的旋转，或者由点样仪或扫描假象造成的对称性的偏离。后两个参数是自动网格定位算法中的重要因素。

定位过程既要准确，以确保后续图像分析各步骤的精确性；又要高效，以使得玻片处理能够高通量，这一点非常重要。

允许用户干预可以提高定位阶段的可靠性，尽管这样做可能会使得处理过程非常缓慢。在关于微阵列的文献中，定位通常称为“网格化”。现在的大部分软件系统都提供了手工和自动的网格化过程。定位过程的变化很多，至今仍没有被证明是很好的方法。

分割

图像分割通常定义为将图像划分为若干不同的区域，在每个区域中有着特定属性的



528

图 7-6 一个典型的微阵列玻片的布局，其中由 4 行和 4 列网格组成大网格，并且每个网格内含有 9 行和 10 列样点

过程（Soille 1999）。在微阵列实验中，分割就是将像素归类为前景（即在样点内部）或者背景。对微阵列图像分割后可以计算每个被点样的 DNA 序列的荧光强度，用于测量微阵列所检验的两个样本的相对转录丰度。每种分割方法产生一个样点模板（mask），由属于每个给定样点的前景像素集合构成。根据其产生的样点几何形状，现有的微阵列图像分割方法可以分成四类（表 7-5）：

- 固定圆分割
- 适应性圆分割
- 适应性形状分割
- 直方图分割

表 7-5 分割方法和算法范例及软件实现

固定圆	ScanAlyze，GenePix，QuantArray
适应性圆	GenePix，Dapple
适应性形状	Spot，区域增长和分水岭
直方图	ImaGene，QuantArray，DeArray，适应性阈值确定

固定圆分割

固定圆分割用一个恒定直径的圆来匹配图像中的所有样点（图 7-7，左）。这种方

法容易实现，而且当图像中所有的样点都是圆形并且大小相同时，这种方法的效果很好。该方法首先在 Esien (http://rana.lbl.gov/EisenSoftware.htm)编写的 ScanAlyze 软件中实现，并且成为大多数软件的一个标准选项。然而，对于样点大小有变化的微阵列，固定直径分割的效果并不令人满意。

适应性圆分割

在适应性圆分割中，对于每个样点单独确定用来定义前景的圆的直径（图 7-7，中）。在 Axon 扫描仪专用的软件 GenePix 中实现了这一类算法（Axon Instruments http://axon.com/GN_Genomics.htm/#Software）。注意在 ScanAlyze 和其他类似的软件包中，提供了可以让用户手工逐点调整圆的直径的选项。在实际应用中，这样做会非常耗费时间，因为在每个微阵列中含有数千个样点。适应性圆分割的另一个例子是 Dapple 软件（Buhler et al. 2000)，该软件通过确定样点的边缘来寻找样点。简而言之，Dapple 软件计算图像的负二阶导数（拉普拉斯算子，Laplacian）。在 Laplacian 图像中数值高的像素对应于样点的边缘。此外，Dapple 软件通过在 Laplacian 图像中寻找最亮的环（圆）而给样点强加上圆形的约束。

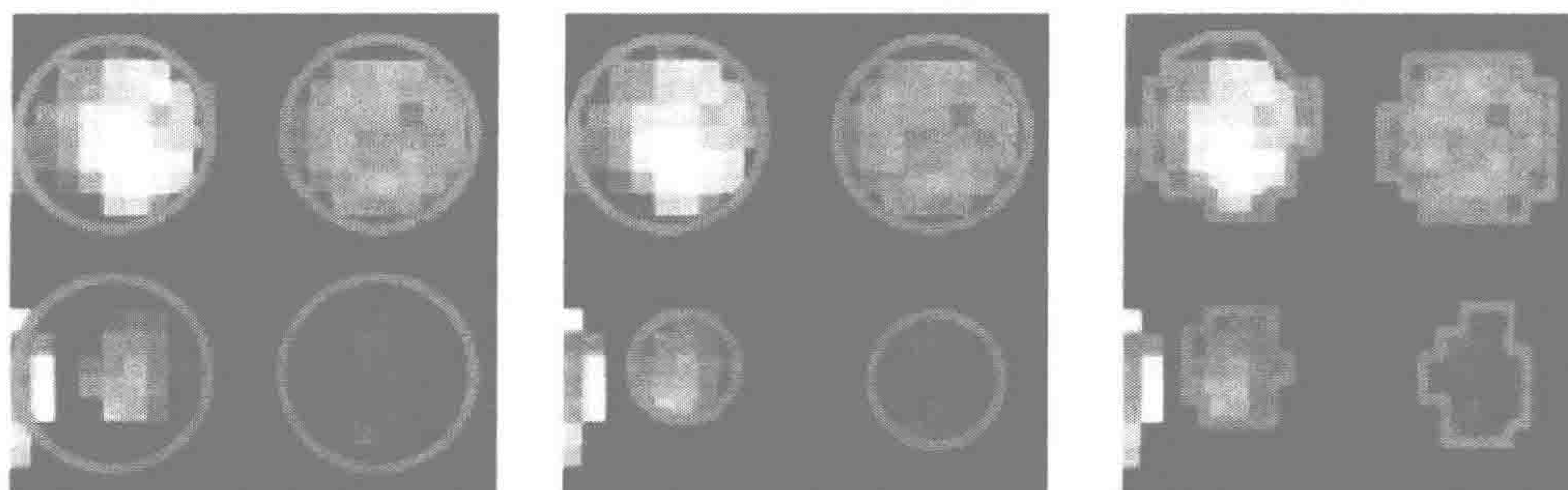


图 7-7 固定圆（左）、适应性圆（中）以及适应性形状（右）分割的分割效果示意图。全部三种分割过程均生成一个样点模板，位于其中的像素被认为是前景，也就是说，用作基因表达水平的测量（见彩图）

529

适应性圆分割方法对于制作良好的微阵列而言效果非常好，但是对于椭圆形或者圆环形状的样点效果不好。非圆形样点通常来源于点样过程（如点样针尖的形状、溶质的不均匀沉积）和点样后的玻片后处理效应（Eisen and Brown 1999)。因此更加希望对样点的形状没有限制的分割算法。

适应性形状分割

图像分析中两种常用的适应性形状分割（图 7-7，右）是分水岭法（watershed）(Vincent and Soille 1991; Beucher and Meyer 1993）和种子区域增长法（seeded region growing，SRG)（图 7-8)（Adams and Bischof 1994)。在本书写作时，适应性形状分割方法还没用集成到应用最广泛的微阵列分析软件包当中。

分水岭和 SRG 分割都要求预先确定起始点或种子，使用这些方法的图像分割方式

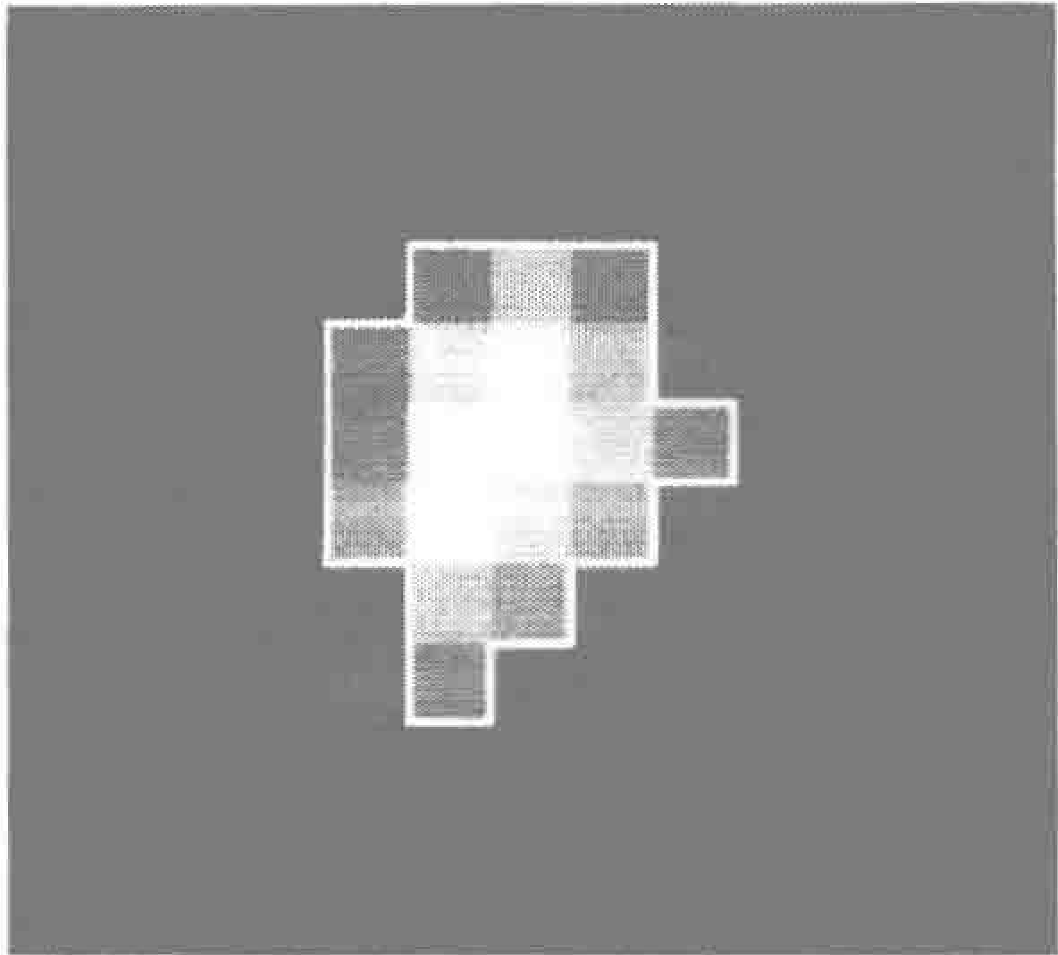


图 7-8　图中白色粗线表示 SRG（种子区域增长）对非圆形样点进行图像分割的结果

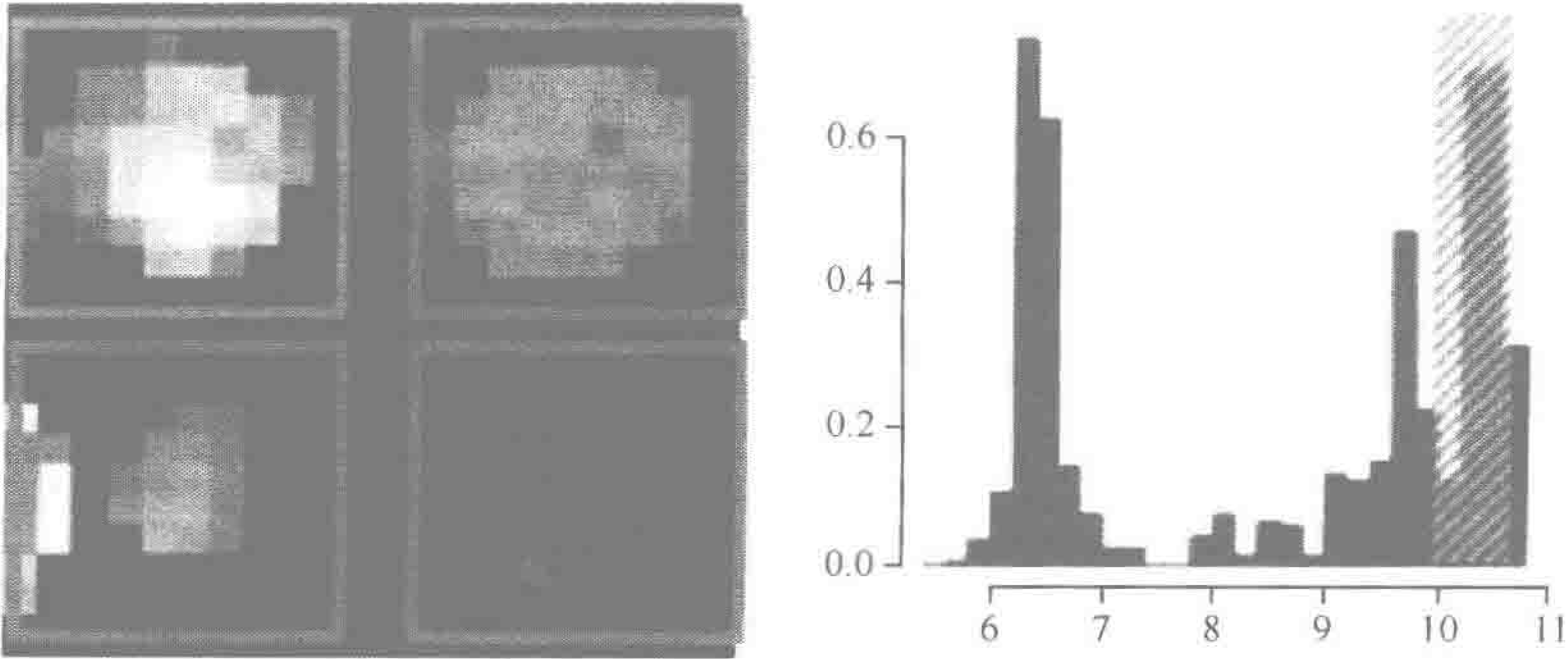
白色粗线内部的像素归类为前景，其余的像素归类为背景（见彩图）

530

的弱点主要是种子点的数目和位置的选择。然而，在微阵列图像分析中，特征单元（样点）的数目预先是知道的，并且在寻址阶段也已经确定了样点中心的大致位置。因此微阵列图像非常适合于应用这类方法进行处理。SRG 算法已经在 Spot（Buckley 2000，http://www.cmis.csiro.au/iap/Spot/Spotmanual.htm）和 AlphatArray（Wang 未发表）软件中实现。这些程序的优点在于可以处理偏离圆形的样点形状。

直方图分割

直方图分割方法并不清晰地将像素归类为前景或背景。相反，这类方法对指定区域内的像素进行分析，并且根据像素的分布来估记前景的强度。这类方法应用一个比任何样点都要大的目标模板（target mask）。对于每个样点，用处于模板区域内部像素的像素值的直方图去估计前景和背景强度。因此这类方法没有利用任何的局部空间信息。采用直方图分割方法的一个例子是在 QuantArray 软件中所实现的，其中前景和背景定义为平均强度位于某些预设的百分数范围内。背景的缺省百分数设置为 5％和 20％，前景的缺省百分数设置为 80％和 95％。图 7-9 显示了青色方框区域内像素的分布，阴影区域覆盖了落入 80％和 95％百分数之间的像素值。另一个实现直方图分割方法的例子由 Chen 等（1997）描述。他们使用了一个圆形的目标模板，并根据 Mann-whitney 检验来计算出阈值。大于阈值的像素归类为前景，否则为背景。该方法在用于 GSI Lumonics 扫描仪的 QuantArray 软件和 Scanalytics 公司的 DeArray 软件中得到实现。直方图方法的主要优点是简便。然而，它的一个主要缺点是在设定一个大的目标模板来补偿样点大小的变化时，定量的结果不稳定。



531

图 7-9　样点描绘的直方图方法

图示的在 QuantArray 中使用的直方图方法。包围每个样点的正方形表示了目标模板。直方图中表示的是青色正方形（左上）内部的像素分布。直方图的 x 轴表示像素强度的对数变换。缺省情况下，估计的前景强度为位于 80％～95％（阴影区域）像素值的平均。而且，直方图分割产生的样点模板不必是连通的（见彩图）

图 7-10 中显示了 SRG 和直方图分割法在处理三类不规则样点时的结果比较。第一类是所谓的鬼点，样点的背景值比前景值高。第二类常常称为彗星状点，样点信号受到实验中假象的污染，它们的前景和背景像素强度常常显示出非常小的差异。第三类样点称为圆环形点。圆环形点中央部分的像素强度值要比周围像素的强度值低。这些不同类型的样点给分割算法提出了挑战。图 7-10 中显示了对于全部三类样点，SRG 可以满意地识别出连通的“样点状”区域，尽管对于彗星状点，该区域可能会太大，并且包含了来自 DNA 覆盖区域以外部分的信号。与之相比，直方图方法分别挑选出具有合适的低和高强度的像素作为背景和前景，但是这些像素没有连通成样点状的区域。因此，对于

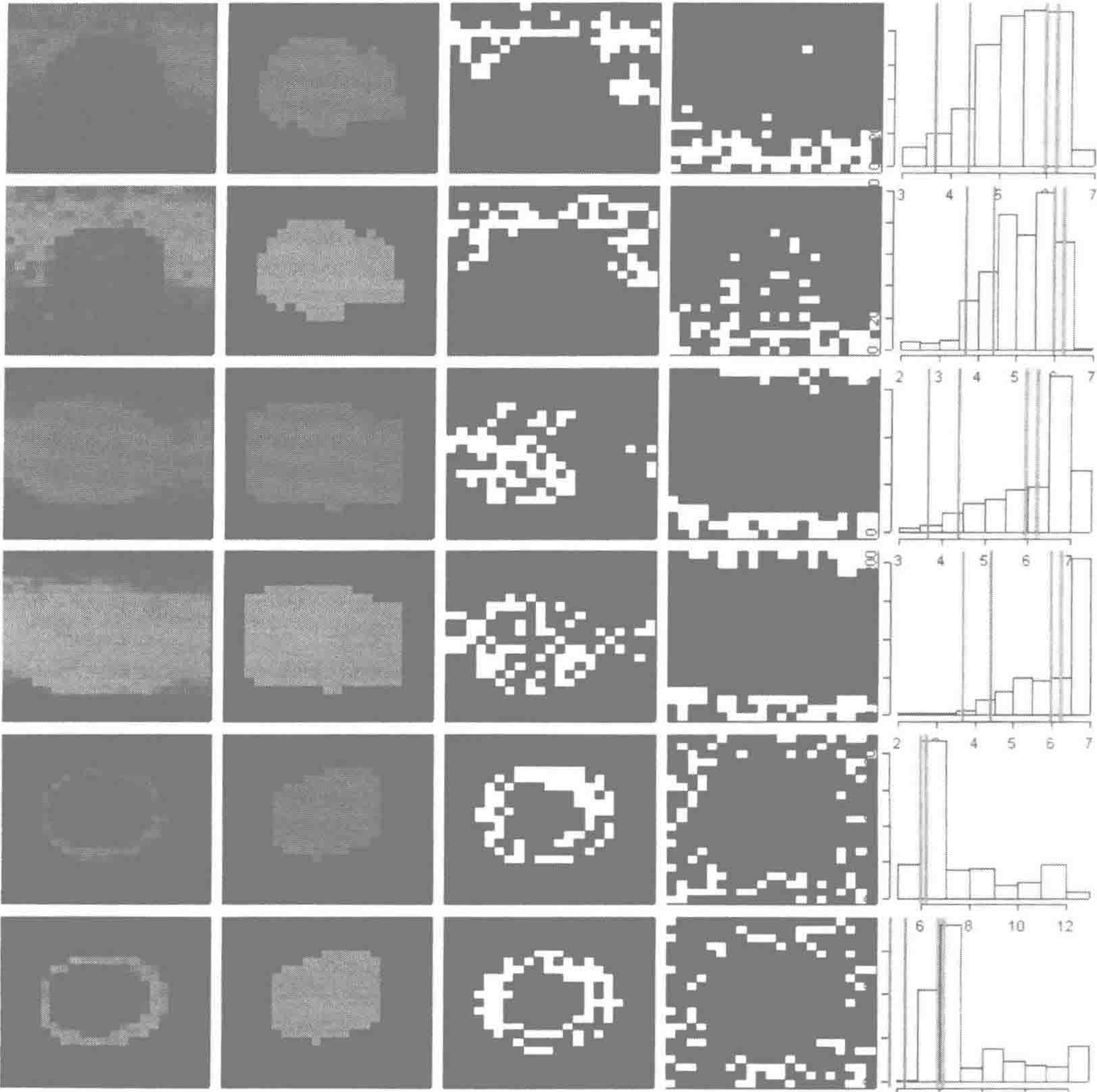


图 7-10 使用 SRG 在图像质量差的区域中识别样点

图中成对显示了常见的三种质量差的样点，分别代表红色和绿色通道。第一列是样点的彩色图，第二列代表使用种子区域增长（SRG）算法得到的样点模板。第三和第四列中白色像素分别代表直方图方法估计的前景和背景中所包含的像素。第五列是第一列图像中像素分布的直方图。在第五列中表示了前景和背景的直方图，以及所使用的像素值的范围（见彩图）

532

彗星状点和圆环形点通过计算前景像素强度的总和来估算样点强度也许是合理的，但是对于鬼点来说理由就不是很充分了。

强度提取

样点强度

在扫描后的图像中，每个像素值都代表了玻片上特定位置处的杂交水平。某个特定的点样后的 DNA 序列的总体杂交与该样点总的荧光成正比。因此样点强度的度量很自然是使用样点模板中各像素强度的累加。后面的计算是基于荧光强度的比值，由于均值的比值等于累加的比值，因此计算出样点模板内的平均像素值。同样地，也可以用中值的比值来替代，计算出样点模板内像素值的中值。

背景强度

微阵列图像分析中背景的调整非常重要，因为测量每个样点得到的强度中都包含有非特异性杂交和玻片上其他化学物质发射的荧光。除了基于直方图的方法以外，上述的分割过程都为背景计算确定了局部背景区域。

533 软件包中实现的各种背景确定方法可以分成以下四类。

局部背景强度

这类背景强度通过将注意力集中于包围在样点模板周围的小区域进行估算。通常，对背景的估计是这些特定区域内像素值的中值。目前大部分的软件包使用这种方法。图 7-11 示例了不同的局部背景调整方法。一种方法（应用在 ScanAlyze 中）是包括以样点为中心的正方形区域内的所有像素并剔除那些在样点模板内部的像素（图 7-11，蓝色方框）。

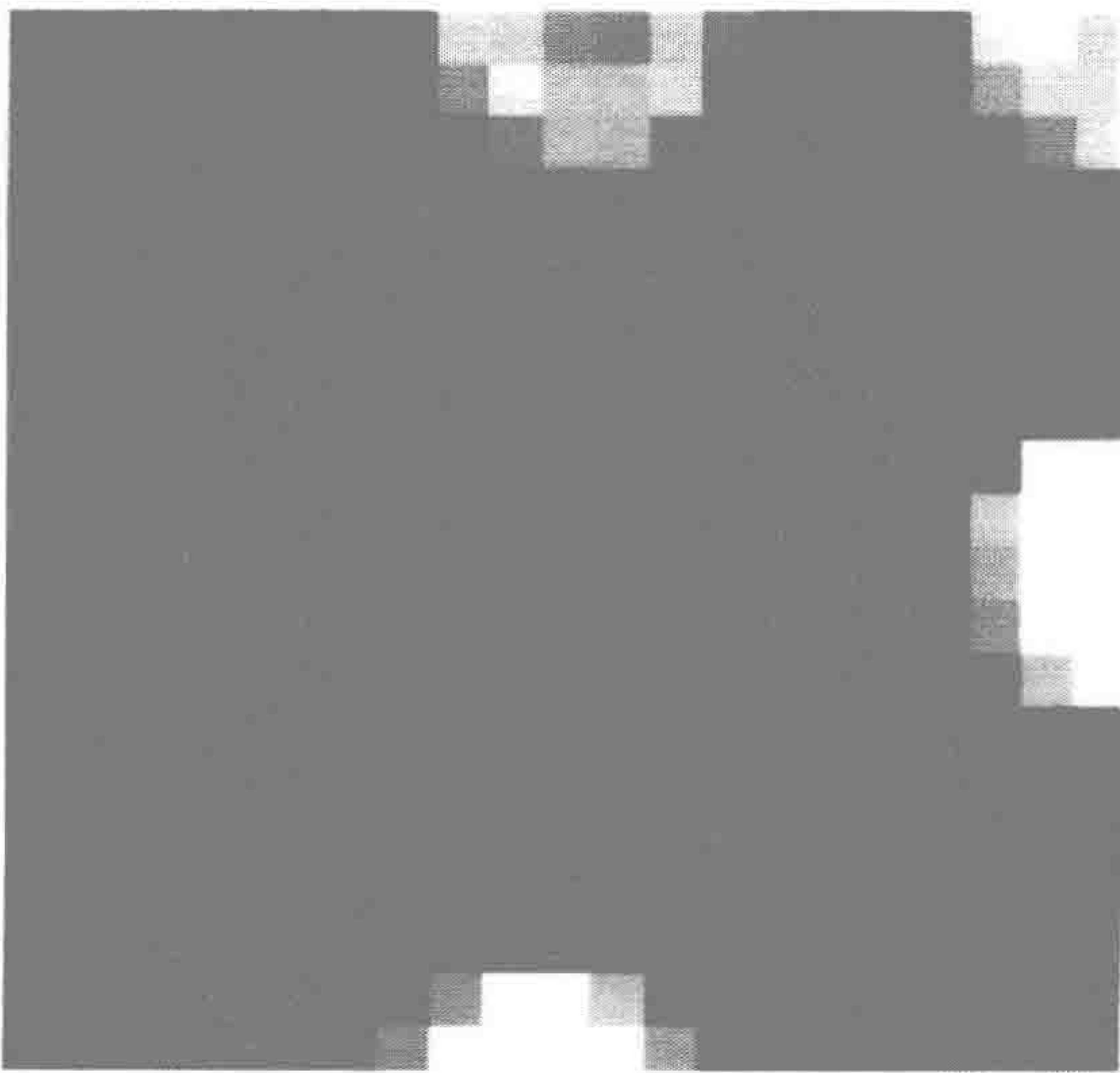


图 7-11　不同局部背景调整方法的图例

红色圆圈内部的区域表示样点模板，其他用彩色线包围的区域分别代表了在不同方法下用于局部背景计算的区域。绿色应用在 QuantArray 中；蓝色应用在 ScanAlyze 中；粉红色应用在 Spot 中（经许可复制自 Yang et al. 2002a）（见彩图）

在QuantArray和ArrayVision中应用的另一种方法是对位于两个同心圆之间的区域进行估计（图7-11，绿色圆圈）。由于不是考虑直接包围在样点周围的像素，因此估计出的背景强度对分割过程不敏感。最后，局部背景也可以在与包围在其周围全部四个样点距离最远的区域内进行估计，也称为“谷”区（图7-11，粉红色菱形区域）。每个样点的局部背景强度用该样点周围四个“谷”区内的像素的中值来估计。根据软件的不同，局部“谷”区的形状各不相同，但是这种背景强度的估计方法在一定程度上独立于分割的结果。GenePix中使用该方法进行背景估计。使用远离于全部样点的谷区像素，在很大程度上确保了背景估计不受到属于样点的像素的破坏。在其他方法，特别是ScanAlyze的方法中，可能会出现因为明亮像素的破坏而使得背景估计偏高。使用远处的像素减小了这种偏差效应，但是必须使用较少数目的像素，并由此造成估计的方差升高，这是对偏差-方差进行折中的一个例子。

形态开

第二种背景调整方法依赖于一种称作形态开的非线性滤波器。使用相同的窗口，先计算局部最小滤波（腐蚀），然后是局部最大滤波（膨胀）而实现该滤波。

在微阵列图像中，使用一个比任何样点尺寸都要大的窗口，其效果是将所有的样点用邻近的背景值代替（有关这类滤波器的详细介绍，见Soille 1999）。在Spot软件中，用一个边长至少为样点间距两倍的正方形结构单元对原始图像进行形态开滤波。该操作除去了所有的样点，并产生一幅作为对整张玻片背景的估计的图像。对于单个的样点，针对该背景图像在名义上的样点中心进行取样来估计每个样点的背景。因为使用了大的窗口来生成形态背景图像，所以预期的空间改变小。形态开与其他简单一些的方法相比较，得到的背景估计要小。然而更为重要的是，形态背景估计的变异性预期要比其他方法低，因为对样点背景的估计是基于一个大的局部窗口内的像素值，然而又不受到属于样点或是处在边缘上的明亮像素的影响而偏高。

534

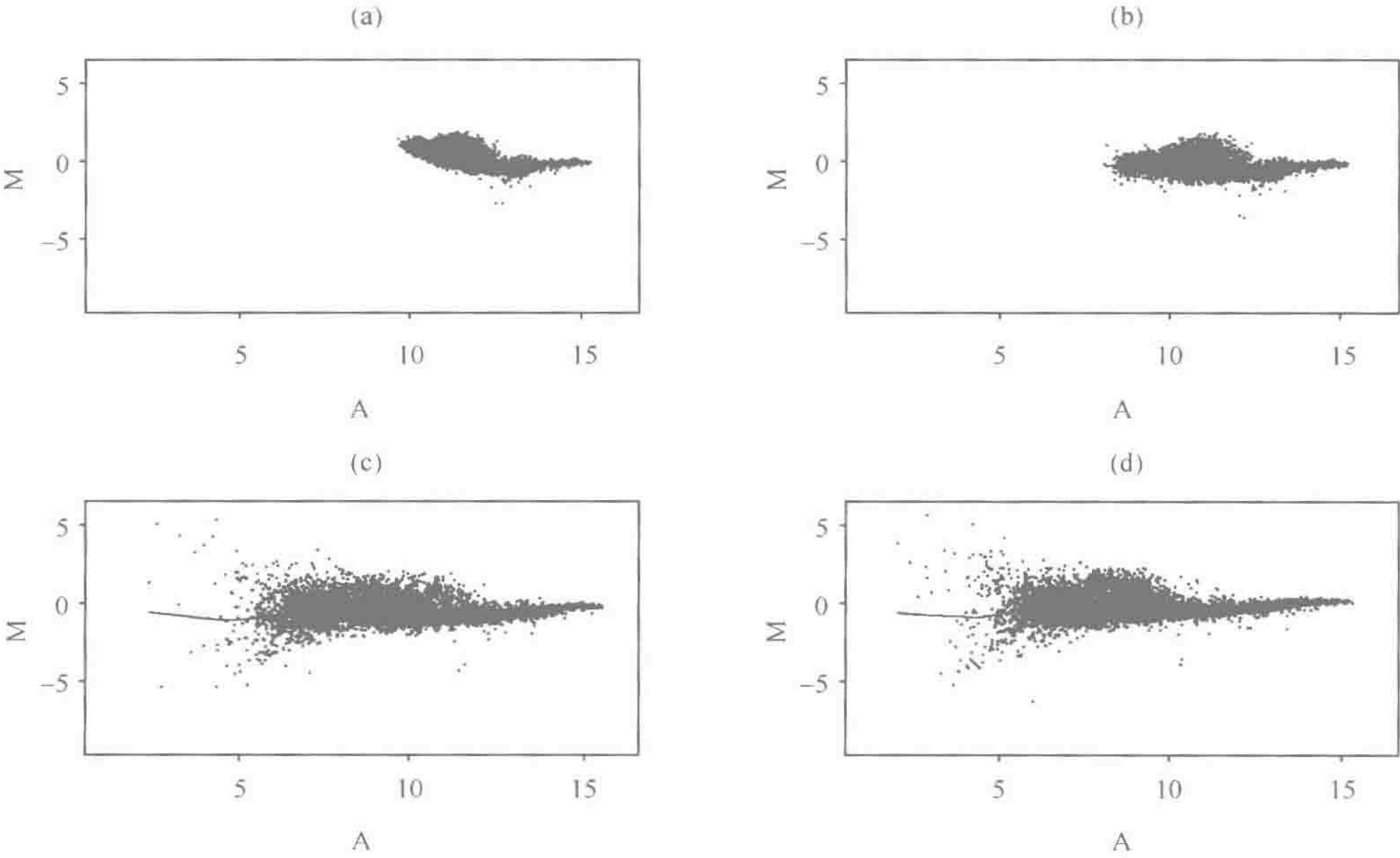
恒定背景

有另外一种估计噪声的方法，该全局性方法从全部样点中减去一个恒定的背景。前面介绍的各种方法均假定对样点信号的非特异性影响可以根据样点周围的区域来估计。然而，一些研究（Lou　未发表结果）表明，荧光染料与阴性对照样点的结合，如对应于植物基因而不会和人的mRNA样品发生杂交的那些样点，要比与玻璃载玻片的结合低。如果情况是这样的，根据一组阴性对照样点的集合来估计背景会更加有意义。如果在整张玻片上背景的分布不均匀，如由于洗涤不够充分，此时这种方法的价值就变得有限。

无背景调整

最后，我们也来考虑根本不对背景进行调整的可能性。对不同方法的比较（Yang et al. 2000）发现，选择背景调整方法要比所使用的分割方法对强度的对数比值的影响更大。因此，寻找更好的分割方法不如选择一种稳定的背景调整方法那么重要。在对背景影响的估计方面，Yang等（2000）指出，与其他方法相比，形态开对背景的估计要

更好一些。形态背景校正后，在玻片内部以及玻片与玻片之间，计算出的对数比值 $\log_2(R/G)$ 往往表现出较低的变异性。不仅如此，该方法看来不会影响准确性。图 7-12 中显示了四幅 MA 作图，其中用强度比值的对数 $M=\log_2(R/G)$ 对平均对数强度 $A=\log_2\sqrt{R \cdot G}$ 进行作图。图 7-12 中不同的图版显示了同样一幅图像经过不同图像分析方法定量后的 MA 作图。良好的图像分析方法应该能够清楚地将差异表达的基因与噪声区分开来。注意到与不使用任何背景扣除方法（图 7-12a）或者使用形态背景调整方法（图 7-12b）相比较，局部背景调整方法（图 7-12c 和 d）在低强度点的附近显示出较大的变异性。局部背景调整往往会模糊差异表达的基因与噪声之间的区别。



535

图 7-12　四幅 MA 散点图，其中强度比值的对数 $M=\log_2(R/G)$ 对平均对数强度 $A=\log_2\sqrt{R \cdot G}$ 进行作图

(a) 没有扣除背景的样点；(b) 用形态开进行背景调整的样点；(c) GenePix；(d) ScanAlyze.（经许可，复制自 Yang et al. 2001a）

对不同的背景校正方法的比较表明，基于局部领域内的均值或中值所作的估计往往噪声会很大，并有可能导致对数比值的标准差加倍。在另一个极端，不进行背景调整会降低鉴别差异表达基因的能力（Yang et al. 2001b）。

目前尚未出现一种被广泛接受的校正样点强度当中背景的方法。看来在整张玻片上应用大量的阴性对照样点似乎可以得到较好的背景调整，它将不同于像上面讨论的那样，通过对玻片的非样点区域内的荧光强度进行测量而得到背景。任何建议都是暂时性的，但是此刻，我们的建议是用一种比局部背景方法使背景的改变更为缓慢的方法去调整背景。

PART Ⅲ　标准化

536

Yee Hwa Yang, Terry Speed（加州大学统计系，伯克利，加利福尼亚 94720）

导言

微阵列图像分析提供了关于红色和绿色通道中前景和背景的测量值，以及阵列中每个样点的一系列质量度量。使用该信息的下一个步骤是对每个通道相对强度中的系统性偏差进行调整，这一过程称之为标准化（normalization）。标准化的目的还在于校正同一块玻片上的样品之间，或者是不同玻片之间其他的强度差异，这些差异不能反映样品间真实的生物学变化。

标准化的必要性在自我比较实验中最为明显，这时使用不同的染料对两份相同的 mRNA 样品进行标记，并在同一张玻片上进行杂交（图 7-13）。因为在自我比较实验中并没有真正的差异性表达，红色和绿色的强度应该是等同的。然而，由于 Cy5 核苷酸的掺入相对较低，红色强度往往要比绿色强度低，而且在整个阵列当中和阵列之间这种红色和绿色强度的不均衡在所有的样点中通常也不是恒定的。此外，全部样点的强度会

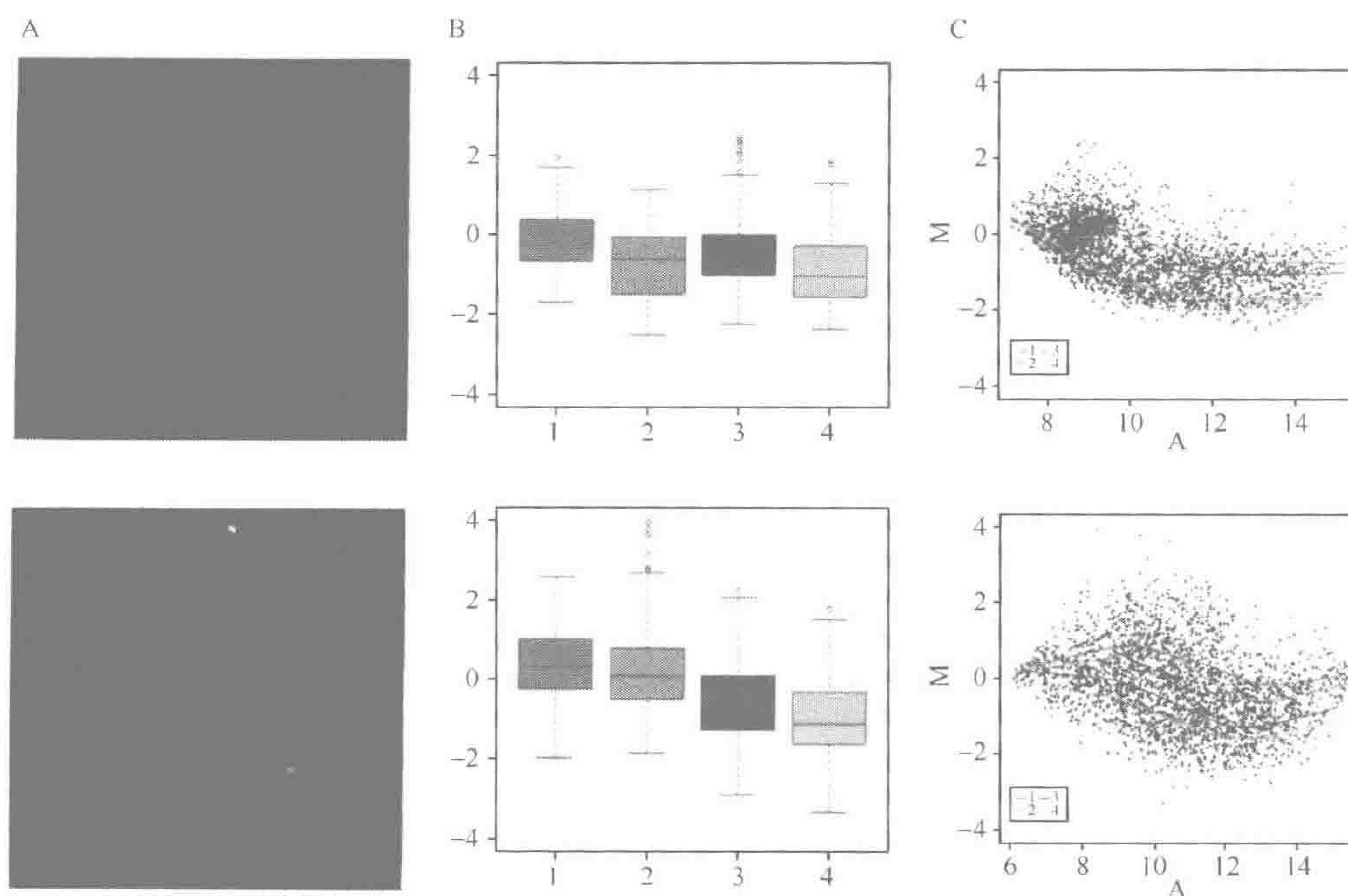


图 7-13　自我比较杂交实验，用于说明在许多微阵列实验中可见的系统性偏差

A. 伪彩叠加图；B. 用箱线图表示从构成 a 中每个阵列的四个方格中得到的数据，注意 $\log_2 R/G$ 比值的平均是方格的函数，因不同方格而异；C. 散点（M-A）图，其中的彩色曲线表示对每个方格中数据最佳的拟合曲线。注意 $\log_2 R/G$ 比值（M）随信号强度 A 而变化。关于使用箱线图和 M-A 作图来表示数据的进一步说明，见 PART Ⅰ。在 PART Ⅲ 中也示例了玻片之间的变异（见彩图）

537

随着在阵列上的位置、点样时加样板的来源、使用的点样针以及玻片扫描参数的不同而变化。造成染料偏差的因素很多，包括染料的物理性质（光和热敏感性、半衰期）、染料的掺入效率、杂交和处理过程中实验条件的改变、每个样品中应用的 RNA 量的差别，以及与特定扫描仪器有关的激发和采集过程。在对数据进行更深入的分析之前，必须对同一块玻片内部和玻片之间的荧光强度进行标准化，以使得假阳性降至最低，从而将两个样品之间真正具有差异性表达的基因识别出来。

玻片内部的标准化

下面讨论了许多应用于每个荧光通道数据的标准化方法。

基于对强度值进行全局性调整的标准化方法

该方法假定整张玻片上的红色和绿色强度之间通过一个常数因子相关联。例如，假设 $R=kG$，并将对数比值分布的中心位置移动到 0：

$$\log_2 R/G \rightarrow \log_2 R/G - c = \log_2 R/(kG)$$

对常数 c，两个常用的选择是表示为 $c=\log_2 k$，或者 $k=2^c$。第一种选择（$c=\log_2 k$）常常称为基于比值统计（ratio statistics）的标准化，并与一组特定基因的对数强度比值的中值或者均值相关联。

该方法通常假定在一个特定的基因子集中，如持家基因，对数比值的均值为零。下面要讨论那些可以应用于该方法的可能基因集。在最早的一个标准化方法当中，Chen 等（1997）采用了一种基于比值统计的迭代方法来估计标准化常数。另一种方法，称为总体强度标准化（total intensity normalization），它通过红色总强度和绿色总强度之间的比值 $k=\sum R_i/\sum G_i$，其中 i 表示特定基因的索引，来估计 $k=2^c$。该方法假定尽管在两个样品之间确实有某些基因发生了差异性的表达，但是在一个典型的微阵列上所表示的数千个基因当中，绝大多数是同等表达的，并且，更进一步，那些上调或者下调的基因相互之间会产生抵消。如果真是这样，每个通道强度值的总和应该相等，当出现差异时，可以计算出一个标准化数值并用来对微阵列上每个基因的强度值进行重新标定。Kerr 等（2000）提出应用 ANOVA 模型对基因表达数据进行分析，并在模型中引入染料偏差来进行标准化处理。Wolfinger 等（2001）也提出一种进行标准化的 ANOVA 模型，其中考虑了处理的效应和微阵列的偏差，然后微阵列上的每一个基因都使用由 ANOVA 模型得到的相同的拟合值来做调整。将这些模型应用于未标准化的数据基本上实现了一个全局性的标准化，但是却没有将在微阵列实验中经常出现的强度或者空间差异，或者其他类型的染料偏差考虑在内。

全局性标准化方法已经被广泛地应用，并在 Genepix 和 GeneSpring 等表达分析软件中实现。在本书写作时，这些方法将玻片作为一个整体来处理，而且正是由于这样，它们既不能校正空间上的差别，也不能处理依赖于强度的偏差（图 7-13c）。

依赖于强度的标准化

红色和绿色通道之间的变化并不总是恒定的，并且可能作为信号强度的函数而变

化。可以用一条通过 MA 作图中间的最佳拟合曲线来校正依赖于强度的变化，并且这条线将成为纵轴新的零线（图 7-13c）（Yang et al. 2002b）。在该分析中，依赖于不同的 A 值，数据对（A，M）中 M 值变动的量为 $c=c(A)$

$$M = \log_2 R/G \rightarrow \log_2 R/G - c(A) = \log_2 R/[k(A)G]$$

对 $c(A)$ 的一种估计是使用 R 软件包（Ihaka and Gentleman，1996）中的 LOWESS（locally weighted scatterplot smoothing，局部加权散点图平滑）函数（Cleveland 1979）对 MA 作图进行局部的散点图平滑处理。散点图平滑方法是一种回归分析，使用从微阵列上未加挑选的基因集合中得到的数据，通过沿着 A 轴方向计算移动平均来进行稳健的局部线性拟合。在这里，稳健指的是曲线不会受到在 MA 图中被看做是局外点的那些少量乃至中等百分比的差异表达基因的影响。由用户定义的参数 f 是在每一点上用于平滑的数据所占的份数，f 越大，拟合就越平滑。通常使用的 f 值为 30%。重要的是，该方法本身假设了：①相对少数的基因发生了差异性表达；②在差异性基因表达和样点本身的强度或位置之间没有系统性关联。在样本之间没有差异性表达的指定基因子集或者对照（见下文）同样可以使用于 LOWESS 平滑。在对具有广泛差异基因表达模式的样本进行比较时，这个方法可能是最为适合的。

还有一些其他的强度依赖的方法被用于局部标准化。Finkelstein 等（2000）给出了一种迭代的线性标准化方法，也称为稳健线性回归。这种线性标准化可以看成是上文中强度依赖的标准化方法的一个更加约束的版本。Kepler 等（2000）提出了一种更为一般的强度依赖的标准化方法，该方法使用了一种与 LOWESS 不同的局部回归方法，尽管也是非常的类似。

实现各种平滑方法的软件可以从 http://cm. bell-labs. com/cm/ms/departments/sia/wsc/ smoothsoft. html 得到，一些商业化的微阵列软件包，如 GeneSpring 也加入了 LOWESS 标准化。公开发布的软件包，包括 SMA（statistics for microarray analysis）（Yang et al. 2001b）和 NOMAD（normalization of microarray data）（Colantuoni et al. 准备中，详见 http://persnerlab. kennedykrieger. org/snomad. htm），也都加入了强度依赖的标准化。这些软件都是基于免费的统计软件 R，它可以从 http://www. r-project. org/下载。

在分属于每个点样针的一组点内部进行标准化

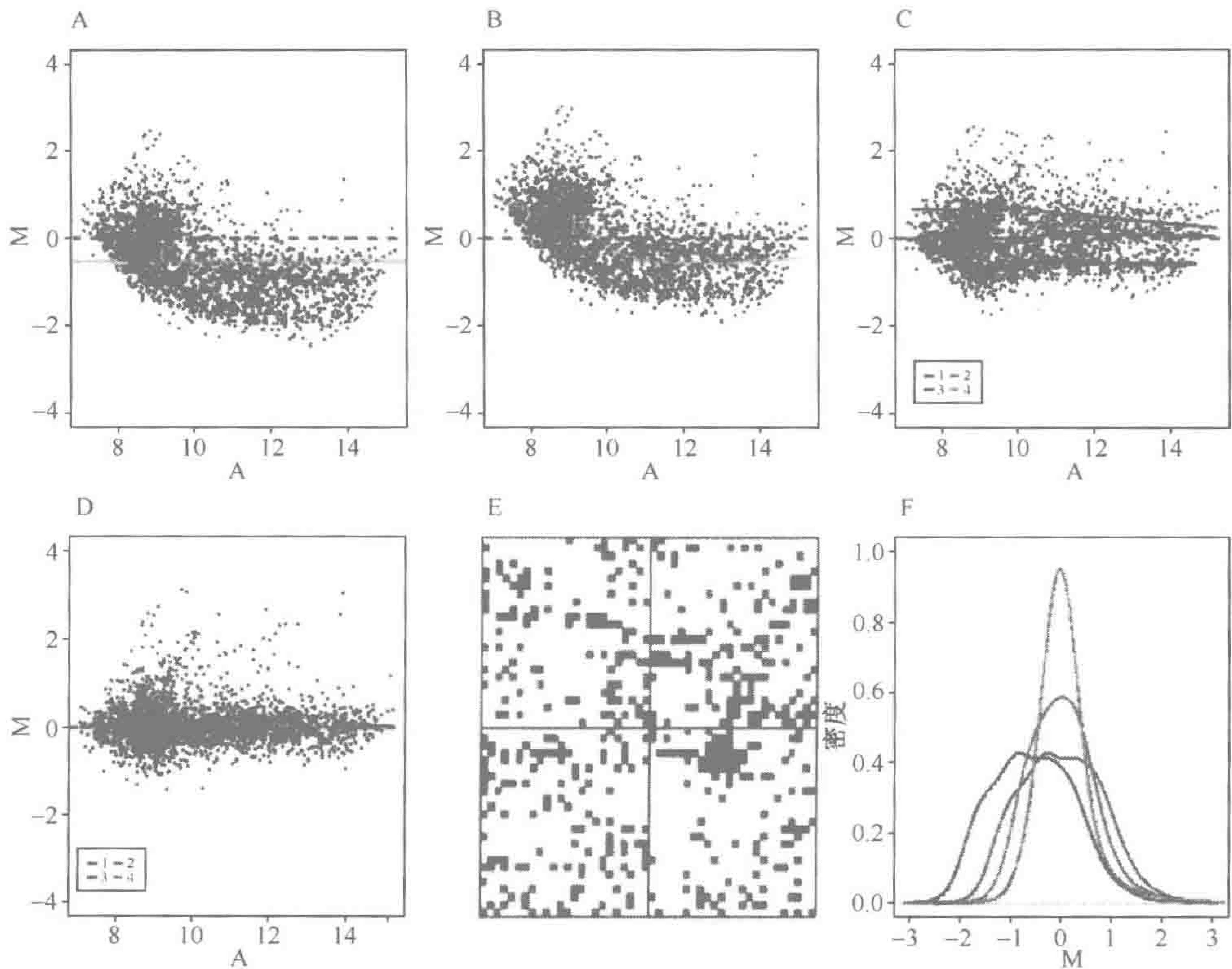
除了强度依赖的变化以外，空间偏差也是系统误差的一个重要来源。大多数标准化方法没有对杂交假象，或者微阵列制备过程中点样针，或是加样板的影响所造成的空间影响进行校正（关于例子，见图 7-14）。图 7-14 中所示的微阵列使用一个 2×2 的点样头进行点样，因此每一条 LOWESS 线对应于由一根特定的点样针生成的一组点（分属于每根点样针的一组点）。从图中可以清楚地看出四条曲线之间的差别是非常大的（图 7-14C）。导致这类变化的来源并不总是很明显，可能是来自于每根针尖的磨损量不同。对于分属于每根点样针的一组点，可以在组内通过 LOWESS 拟合对点样针以及强度依赖的偏差进行校正

$$\log_2 R/G \rightarrow \log_2 R/G - c_i(A) = \log_2 R/[k_i(A)G]$$

式中，$c_i(A)$ 是对 MA 作图中第 i 个方格的 LOWESS 拟合，$i=1,\cdots,I$。I 代表点样针的

数量。

从图 7-14 中可以很清楚地看出标准化对实验精度的改善程度。由于这些数据是来源于自我比较的杂交实验，在图 7-14F 的密度作图中，不同的分散性很好地反映了标准化之后所达到的不同精度。要证实一个基因发生了差异表达，相应的对数比值（M 值）必须要从噪声中显现出来，同时图中这些不同曲线的分散性也说明了经过不同的标准化处理之后噪声的水平。



539

图 7-14　在自我比较的杂交实验中，不同类型的标准化方法对数据分布的影响

A. 标准化之前的散点图。黄色线代表 M 值的中值，它偏离了中值为零情形（蓝色虚线）。B. 使用中值进行标准化后的散点图。黄色线代表通过所有点的全局性 LOWESS 线。C. 全局性 LOWESS 标准化后的散点图。蓝色、红色、青绿色和绿色实线分别代表了通过分属于四个点样针的那些点的 LOWESS 平滑。D. 对分属于每个点样针的每组点进行 LOWESS 标准化后的散点图。E. 对 D 中描述数据的空间作图。F. 在进行任何标准化处理之前（黑色），经过中值标准化后（红色），全局性 LOWESS 标准化后（绿色），以及对分属于每个点样针的每组点进行 LOWESS 标准化后（青绿色）经过平滑处理的 M 值的直方图。全局性中值标准化将对数比值分布的中心移至 0，但是并不会影响到分散性或者是依赖于强度的偏差。强度依赖的标准化与全局性的方法相比减小了对数比值的分散性（见彩图）

玻片之间的尺度标准化

前面的标准化方法目的在于均衡两种染料的强度值，但是并不能使得对数比值的分布一致，例如，在不同的玻片之间。图 7-15 中显示了三个重复玻片在标准化后 M 值的

分散性下降（比较图 7-15 中的 A 和 B）。然而，图中显示不同玻片之间数据的变化程度是不相等的，因为玻片间四分位距（矩形的大小）的差异超出 2 倍以上。这种差别将会导致不同的重复玻片合并在一起时，它们分别有着不相等的地位，也就是说，M 值分散性最大的玻片将对平均值的贡献最大，而这并不是所期望的。作为一个统计学原理，在对同一量的各个估计进行合并时，数据应该经过加权处理，使得其值反比于估计的方差。因此，一种选择是对来自三张玻片的 M 值分别加权处理，而不是直接合并。通过改变来自每张玻片的 M 值的尺度使得每张玻片经过尺度调整后的四分位距相等，也可以实现同样的结果。作者建议，当四分位距的差异超出 2 倍以上，对应于方差的差异为 4 倍或者更高时，应该根据这些方法进行尺度标准化。当从分属于每根点样针的各组点得到的数据的分散性差异显著时，尺度标准化也可以在玻片内部进行（Yang et al. 2002b）。

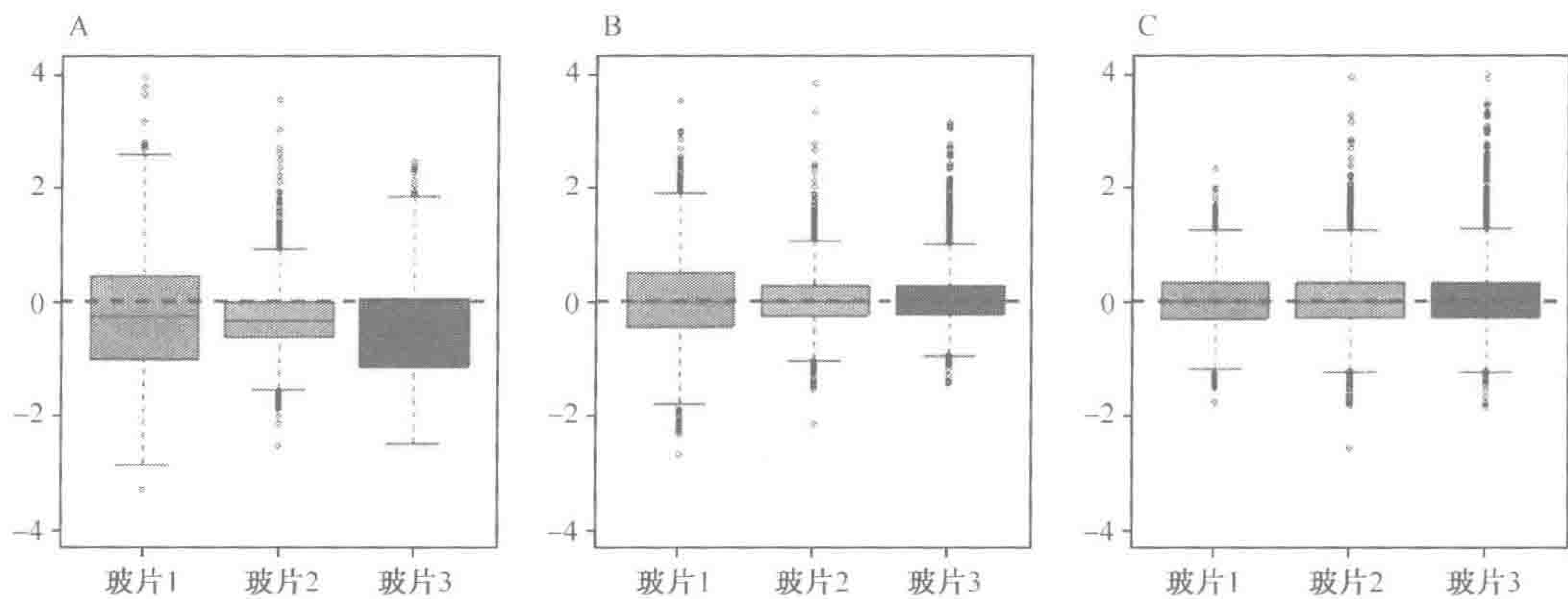


图 7-15　多张玻片的标准化

箱线图表示了在一次自我比较的杂交实验中，三个重复玻片的对数比值（M 值）的分布，标准化之前（A），分属于每个点样针的一组点内部进行标准化后（B）及玻片之间尺度标准化后（C）（见彩图）

使用何种基因集合来进行标准化？

541

对于上面介绍的大多数方法而言，必须决定将采用何种基因集合来进行标准化。通常，最适用于标准化的基因集合取决于实验的性质、所观察到的基因表达变化的量，还可能有对数据标准化所采用的方法。

阵列上的所有基因

微阵列上的生物学比较常常具有非常特殊的性质，并且在基因表达中期望的差异也仅限于一小部分基因。在这种情况下，大量保持不变的基因可以用来表示两个通道的相对强度。这是因为下列的假设是合理的：

- 在两个 mRNA 样本之间，只有相对很小的一部分基因具有显著的表达变化（参见自我比较杂交的特例）。
- 上调和下调基因的表达水平有对称性。

许多标准化方法中可以应用微阵列上的"所有基因"，但是当两个样本之间的表达

模式具有广泛性的差异时，这种方法就不再适合了，因为此时上述的假设不再成立（表7-6）。

表 7-6　各种标准化方法的总结

		玻片内部				玻片之间
		全局 c(・)为常数 a(・)=1	强度依赖 c(・)=c(A) a(・)=1	点样针依赖 c(・)=c(A,针) a(・)=1	点样针尺度 c(・)=c(A,针) a(・)=a(针)	尺度
全基因	假设两个mRNA样本之间的大多数基因具有相似的整体表达水平。	是	是	是	是	是
持家基因	通常有高表达水平并且不具有强度依赖的结构。	是	否	否	否	否
MSP稀释系列	不需要任何预先的生物学假设；但是从一小部分基因估计c(A,针)可能会不很稳定。	是	是	否	否	否
秩不变集	可能无法覆盖全部的强度范围	是	是	否	否	否

行表示不同的标准化方法，列代表所使用的不同基因集合。对于玻片内部的标准化，对数比值用 $\log_2 R/G \rightarrow [\log_2 R/G - c(\cdot)/a(\cdot)]$进行调整，其中c(・)和a(・)对应于局部的和尺度调整，是或否指使用对应的基因集合进行标准化的实用性。例如，只可以使用持家基因进行全局性的标准化，而只有持家基因不能用于强度依赖的标准化。

542　持家基因

持家基因是一组预先确定的、在大量细胞和组织类型中为基本的细胞过程所必需的基因集合，其表达通常不依赖于组织的发育阶段或者生理或病理学状态。虽然肌动蛋白和 *GAPDH* 基因常常被用作持家基因，但是实际上，即使有也是非常少的基因能够符合这个严格的标准。基于持家基因强度值的全局性标准化方法可能会显示出样品特异性的偏差，并且不能解决依赖于强度的染料偏差问题。当指定的持家基因倾向于高度表达，则可能无法代表所感兴趣的低水平表达并且会受到强度依赖偏差影响的基因。

添加对照

该方法包括使用来自非研究生物的 RNA 靶分子或者人工合成的基因，与微阵列玻片上没有真正的同源区域所以不会和杂交探针之间产生交叉杂交。这些"基因"点样在微阵列上，而且它们对应的 RNA 转录物以相同的数量包含在两个靶样本当中。研究人员可以在微阵列上点样一系列包含了不同数量添加对照的样点，并预期这些样点在整个浓度范围内应该显示相同的红色和绿色荧光强度。许多公司（Stratagene 和其他一些）销售含有添加混合物以及持家基因的试剂盒，用于微阵列实验的验证和标准化。

合并的微阵列样品（稀释系列）

Yang 等（2002b）提出使用制备微阵列所用的 cDNA 合并后作为探针并进行一系列的稀释。在理论上，全部 cDNA 序列应该都能够和这个混合探针样品杂交，并且假定总的信号受到任何样品特异性偏差的影响最小。然而，包含表达程度非常高的基因会导致一些序列在来自该混合物的探针信号中的贡献不成比例。因此该信号要从一个复杂性较低的标记靶分子混合物中来获取，而不是像预期的那样。该方法的另一个缺点是标准化值是根据相对数量较少的稀释系列的样点来进行计算的。一个与此相关的方法是使用一系列浓度的基因组 DNA，例如，在酵母基因组中包含有较低程度的非编码序列。

添加对照或者合并微阵列样品（MSP）的稀释系列可以适用于一系列的标准化方法；然而在考虑使用它们时，添加对照或者 MSP 的各个探针在微阵列中的数量是非常重要的。使用"所有基因"进行标准化（见上文阵列上的所有基因）为估计对数比值中强度或者空间依赖的趋势提供了最佳的稳定性，因为数据是根据大量的基因收集而来

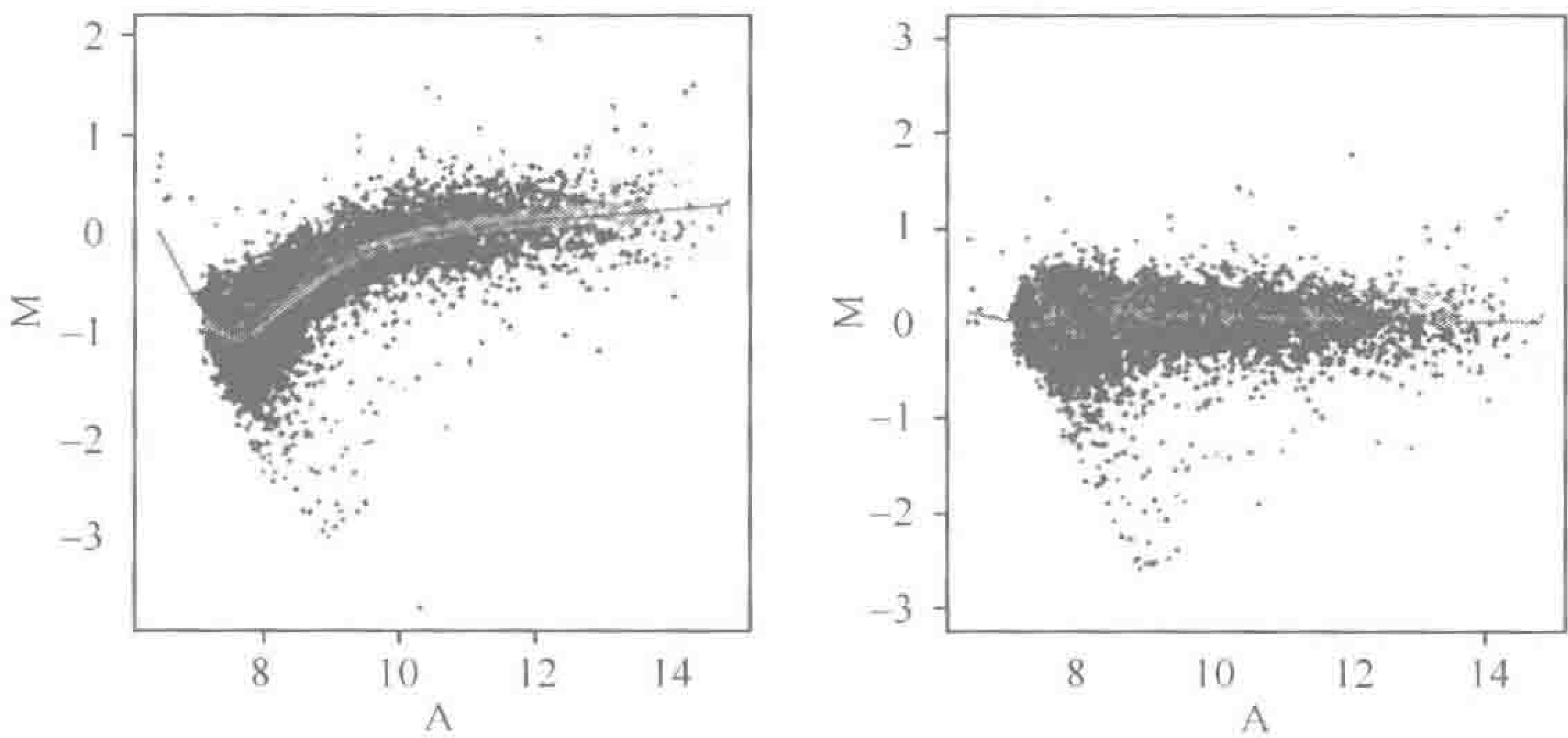


图 7-16 玻片内部的标准化：嗅球前后部分对比的 MA 作图

这些样本非常相似，因此不应期望有许多基因会发生改变。青色点对应 MSP 稀释系列；青色曲线对应的 LOWESS 拟合；红色曲线代表对整个数据集的 LOWESS 拟合。对照基因：（黄色）微管蛋白和 GADPH；（绿色）小鼠基因组 DNA；（橙色）一个大体上的秩不变基因集合，$P=0.01$ 且 $I=25$。（左图）标准化前的 MA 作图。（右图）对分属于每个点样针的每组点进行局部标准化后的 MA 作图。

（经许可复制自 Yang et al. 2002b）（见彩图）

543

的。与此相比，添加对照或者 MSP 稀释系列可能只表示了微阵列上的少数探针，因而比较不适用于空间标准化（Yang et al. 2002b）。当两个样本间的基因表达有大量的差异时，在这种情况下使用“所有基因”进行标准化的价值有限，此时 MSP 或者添加对照可能是最为行之有效的。例如，图 7-16 中显示了当数据来自具有相似的表达模式的样本时，基于 MSP 和“所有基因”的 LOWESS 曲线的一致性，而当样本间具有生物学差异时，曲线就显示出不同（图 7-17）。

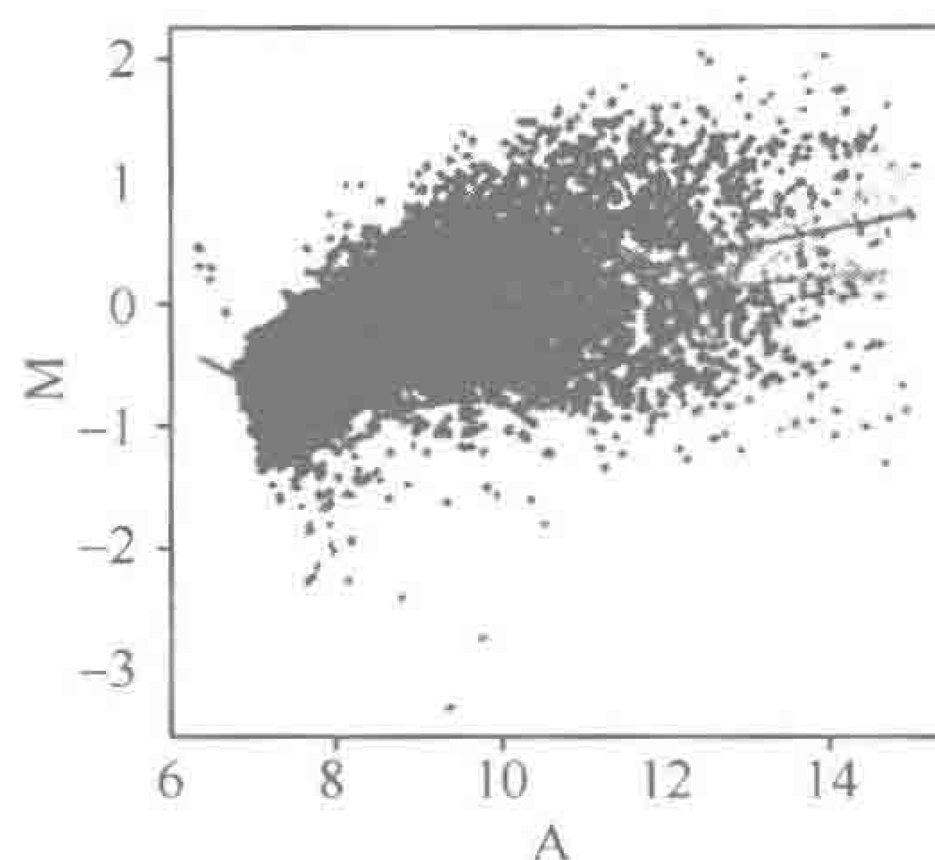


图 7-17　玻片内部的标准化：嗅球中部和侧部对比的 MA 作图

青色点代表 MSP 稀释系列；青色曲线代表对应的 LOWESS 拟合。红色曲线代表对整个数据集的 LOWESS 拟合。绿色曲线表示复合标准化曲线。对照基因：黄色代表微管蛋白和 GADPH；绿色代表小鼠基因组 DNA；橙色代表一个大体上的秩不变基因集合。该图显示了标准化以前的 MA 作图（经许可复制自 Yang et al. 2002b）（见彩图）

Yang 等（2002b）提出了一种复合的标准化方法，对于低强度值的标准化是根据“所有基因”，并且对于那些较高的强度和差异性生物学样本的标准化则是基于 MSP 稀释系列。

其他的基因集合

微阵列标准化也可以使用经过挑选的基因来进行，这是因为在一组特定的实验当中这些基因在所有的样品中是保持不变的。这类方法中的一种是选择一个秩不变的基因集合。如果一个基因集合的红色和绿色强度排序相同，那么就称其为秩不变。也就是说，要使两个基因秩不变，对于一个给定实验的所有样本，在红色和绿色通道中基因 A 的强度值要始终大于基因 B 的强度值。实际应用中，一个最大的不变集对于标准化来说往往还是太小而没有效用。Schadt 等（2000）（也见 http://preprints.stat.ucla.edu/303/）和 Tseng 等（2001）提出了一种迭代的方法来寻找一个大体上不变的基因集合。

总之，尽管在微阵列实验中对人为的偏差进行调整非常重要，理解在每种校正偏差的方法之中包含了对偏差来源的假设也是同样重要的。研究人员应该了解这些假设以使得标准化的算法能够明智地应用，同时应该检查在标准化之后系统误差要下降，观察到的基因表达差异要有意义，而且是真实的。

PART Ⅳ　芯片数据的表示和评价

544

Yee Hwa Yang, Terry Speed（加州大学统计系，伯克利，加利福尼亚，94720）

导言

这部分内容中建议使用多种惯用的图形显示方法来检查微阵列实验的结果。这些显示可以帮助判断实验是否成功，并有助于选择合适的分析工具，以及使特定的实验问题突现出来。因此芯片数据的各种表示能使得将来获得较好的实验结果变得容易。

伪彩色覆盖图

Cy3 的 5′亚磷酰胺结合物的发射峰为 570 nm，Cy5 的结合物为 667 nm，因此扫描后的微阵列输出图像中 Cy3 和 Cy5 通道通常分别表示为绿色和红色的伪彩色。当两幅图像合并在一起时，如果一个给定的基因在两个样品中的表达相同，会得到黄色的点，或者在差异表达时为暗红色或暗绿色。为了进行显示，两幅 16 位 Tiff 图像（来自 Cy3 和 Cy5 通道的扫描输出）通过平方根变换压缩为 8 位图像。采用这种转换是为了使用 24 位复合 RGB 叠加图像来显示两个波长下的荧光强度。在这种 RGB 图像中，蓝色值（B）设置为 0，红色值（R）用来表示 Cy5 的强度，绿色值（G）用于表示 Cy3 的强度。每个通道的协配准和覆盖提供了对微阵列实验快速可视化的一种方法，提供了关于色彩均衡、杂交均一性、样点均一性、背景以及诸如灰尘或划痕之类伪像的信息。覆盖图也提供了关于两个样品之间差异表达的基因数量的大致印象。图 7-18 中显示了两幅质量不同的覆盖图。

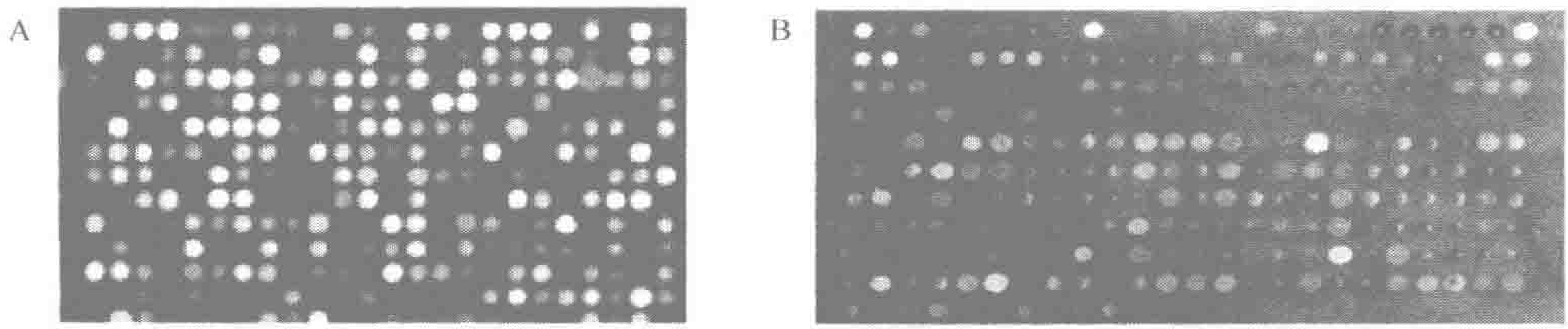


图 7-18　红/绿覆盖图

A. 相比质量较好的芯片，拥有低的背景和较大比例差异表达的基因。B. 一个背景高的微阵列实验，在第三行中有许多鬼点，而且差异表达的基因很少。关于从覆盖图中可以获取的疑难问题判断信息，参见第 3 章，疑难问题解答（见彩图）

覆盖图为微阵列实验提供了有价值的定性印象，但是它们也可能带来误导，特别是当数据通过 Adobe Photoshop 这类程序而被增强时。因此，使用由图像分析软件提取的数据来得到定量的质量评分和检测细微的表达差异是很关键的。

样点强度的对数变换

545

出于多种原因，在作图或者进行深入分析之前将图像分析程序获得的原始强度取对

数是非常有必要的。首先，在一个成功的微阵列实验中，强度通常跨越了整个16位，从0到65 535单位，其中大多数处在数值较低的范围，小于1000。如果数据没有进行变换，在低数值的范围内它们必然要以高度压缩的形式表示出来。进行对数运算可以使数值均匀分布在整个取值范围，并且能提供比较容易的数据可视化。第二，随机变异（由强度的标准差度量）通常随着平均信号强度近似线性地增长。进行对数变换倾向于使得变异更为恒定。第三，$\log_2(R/G)$是差值$\log_2 R-\log_2 G$。将数值表示成以2为底的对数是比较方便的。因此，对数（以2为底的）比值1、2和3分别对应于2、4和8倍的倍数改变。

散点图：对角的和MA作图

散点图是cDNA微阵列实验之后最常用的图形显示方式，它是两个通道强度的$\log_2$作图，$\log_2 R$对$\log_2 G$（图7-19，左）。注意力主要集中在偏离对角线的点，表示了差异表达的基因。数据“云”表示了样点强度的分布。

另一种类型的散点图是MA作图（Dudoit et al. 2002），其中比值信息旋转了45°，而且坐标轴也作了尺度变换。在MA散点图中，以对数比值($M=\log_2 R/G$)为y轴上各点，信号强度几何平均的对数($A=\log_2\sqrt{R\cdot G}$)为x轴上各点进行作图（图7-19，右）。MA作图用来增加表示差异表达范围的空间。之所以使两种强度的几何平均值（A）可见，是因为对转录物丰度整体上的测量为后续的分析提供了一个有用的工具。例如，了解两倍的变化是来自于20 000对10 000的比值，而不是200对100是有帮助的；前者比较不可能是由于随机变化而产生的。

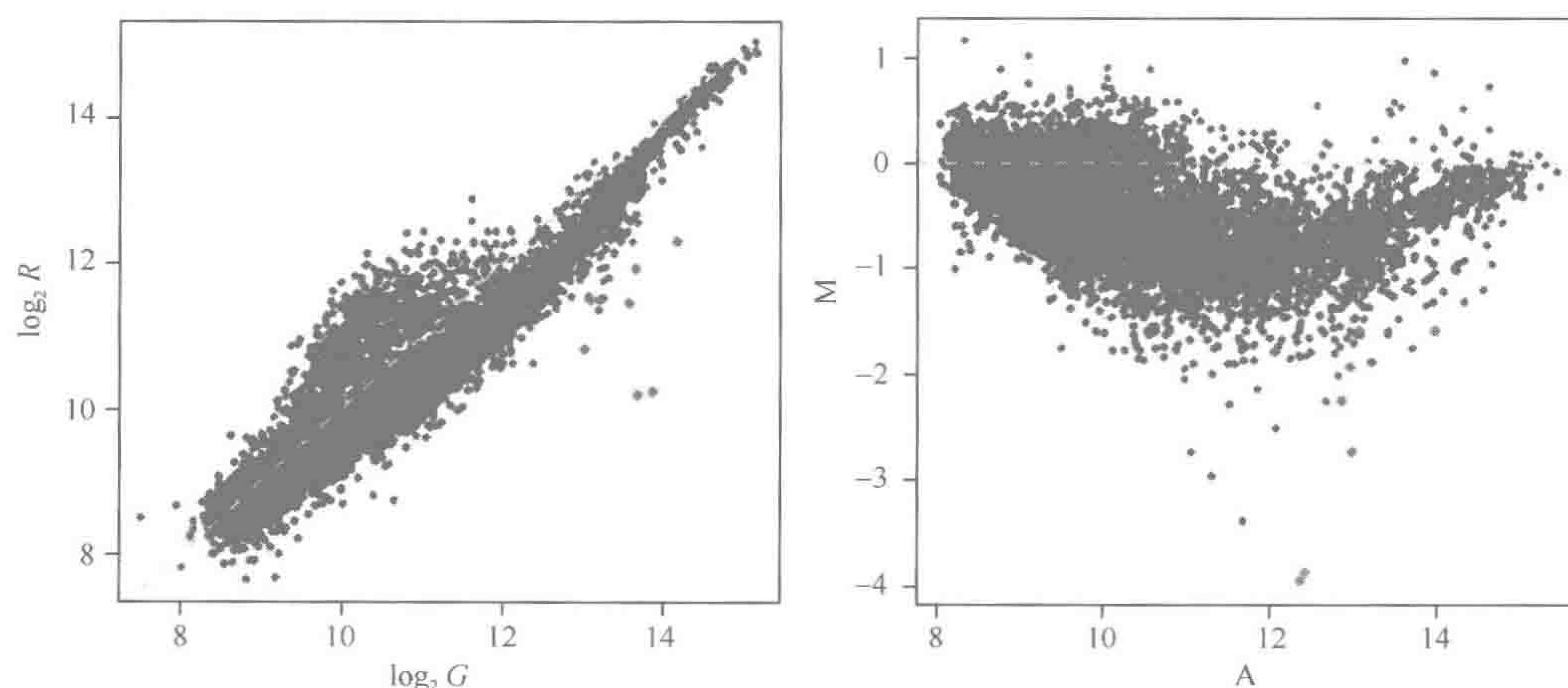


图7-19 log R对log G的散点图（左）和MA作图（右）

（数据来自Callow et al. 2000）。右图中央的下降是一个假象，并且比在左图中更为明显（见彩图）

直方图

直方图提供了关于数据分布形状一个非常有用的快照，也提供了有关定性统计的一些信息。例如，图7-20中显示了每个通道的信噪比（S/N）对两幅不同图像中具有一定信噪比值的样点所占比例的作图。可以看出下方图像中有较高的Cy5背景，因为它

的平均 S/N 比为 2 而上方图像为 4，然而 Cy3 通道却是相当的，具有相近的平均 S/N 比值。

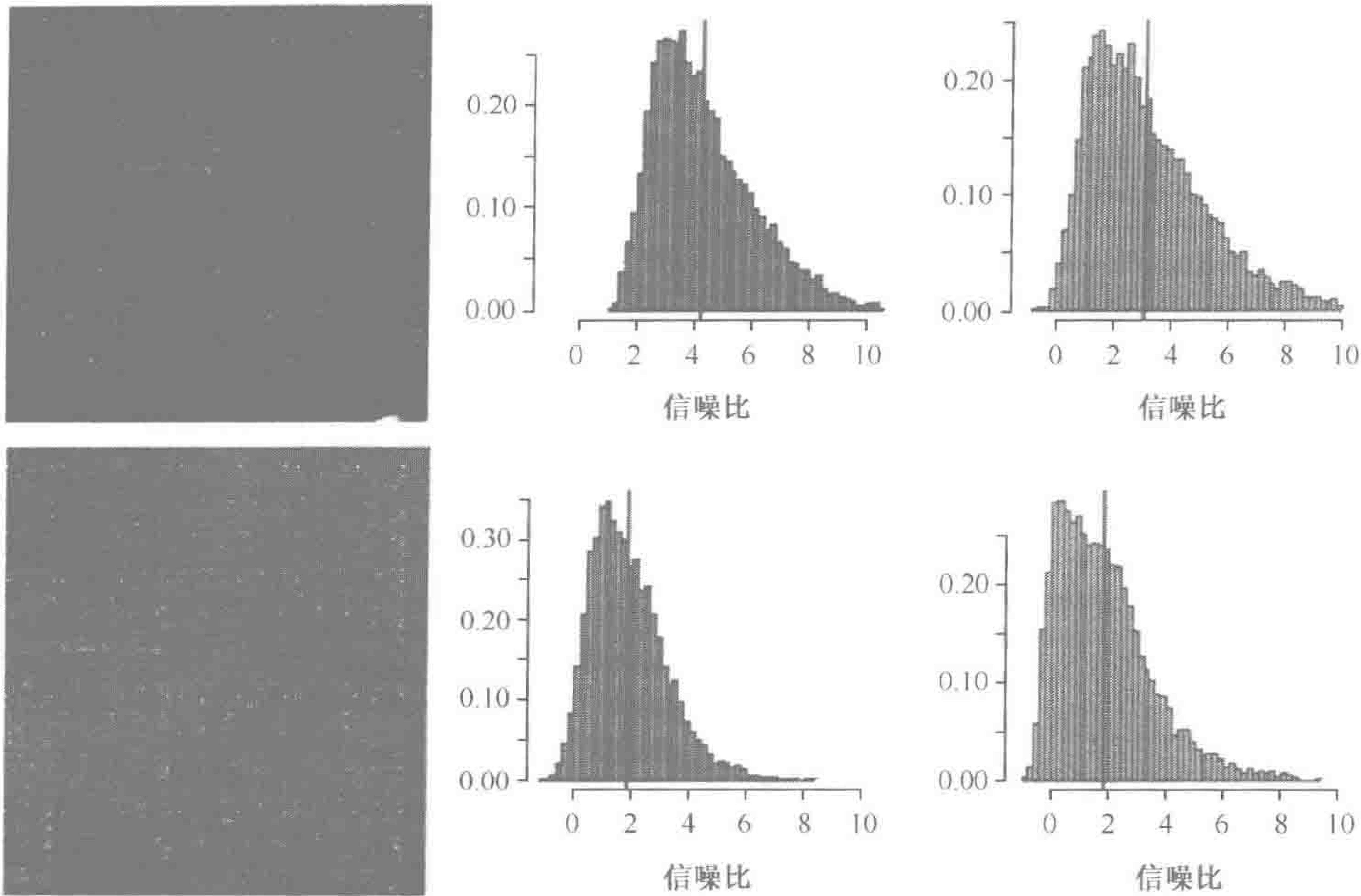


图 7-20 使用直方图在样点水平上评价信噪比（S/N）

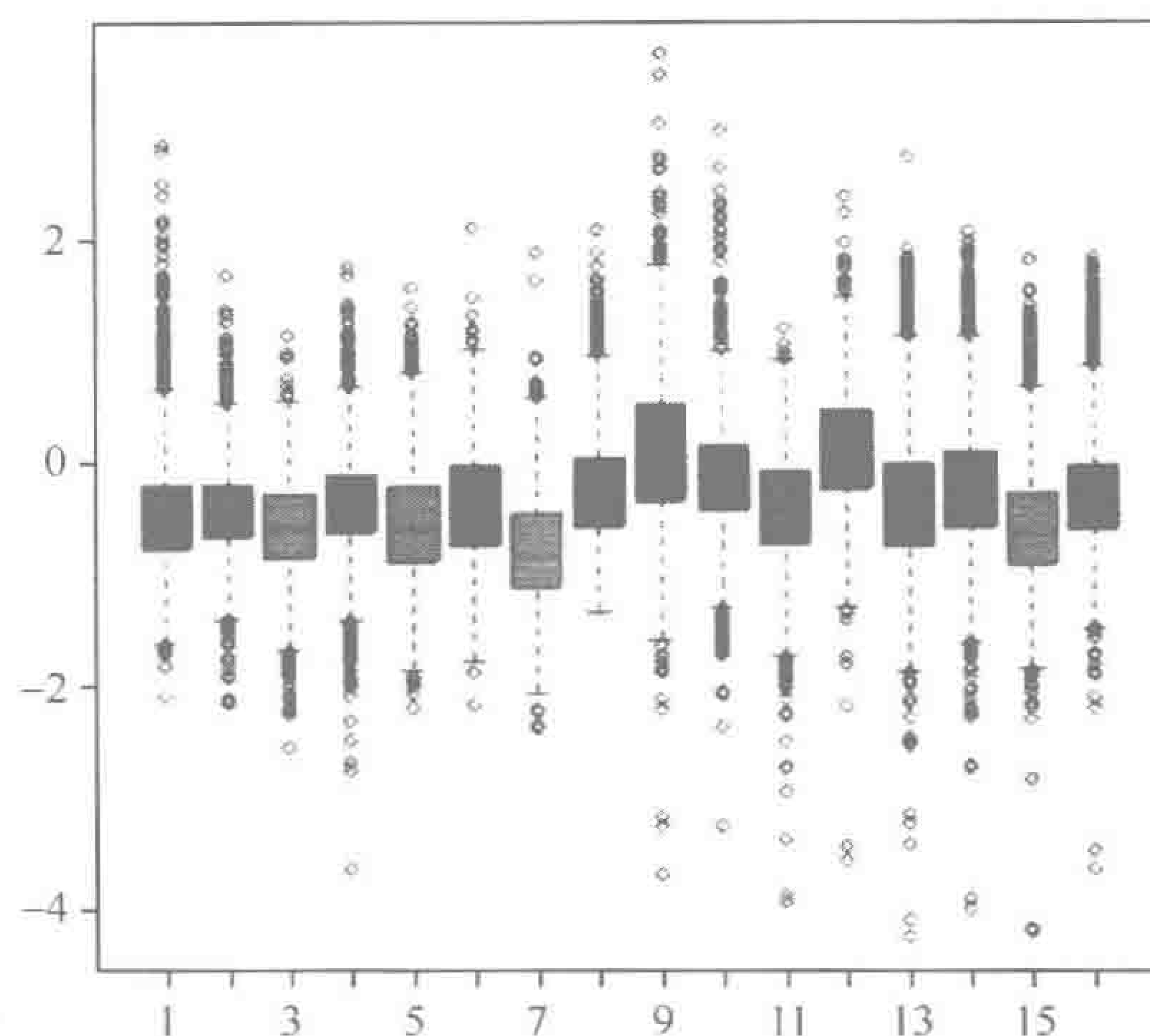
546

信噪比在 x 轴上作图，是 $\log_2$ 度量的（样点强度/邻近样点的背景）。有某个给定值的样点的频率在 y 轴上作图。与上方图像相比，下方的图像在右上角有红色的阴影，这被认为该图像有着较低的平均信噪比。与此相比，Cy3 通道的平均差别较不突出（见彩图）

箱线图

与 MA 作图相配合，对 M 值作箱线图会很有帮助。箱线图（Tukey 1977）也被称作“盒须（box-and-whisker）图”，是对一个数据集合分布的简单图形概括。这种作图显示了由中、上和下四分位数，全距，以及可能的单独极值组成的统计学概要。作图中位于中央的“盒子”表示了四分位距（IQR），定义为第 75 个百分点和第 25 个百分点，或者说上四分位数与下四分位数之间的差值。“盒子”中间的线代表中值，是对数据中间位置的一种度量。高于第 75 个百分点并且＞1.5×IQR 的极值，以及低于第 25 个百分点并且＜1.5×IQR 的极值通常单独作图。具有极端的对数比值的样点表示了可能会有强烈的表达差异的基因。

盒子的大小提供了对实验分散性的一种度量。总希望得到窄的盒子（也就是说，一个较小的 IQR），因为这可以表明实验中的噪声比较少。图 7-21 中表示的是微阵列实验的一个例子，其中显示了使用 IQR 度量的 16 个芯片的数值范围和分散程度。



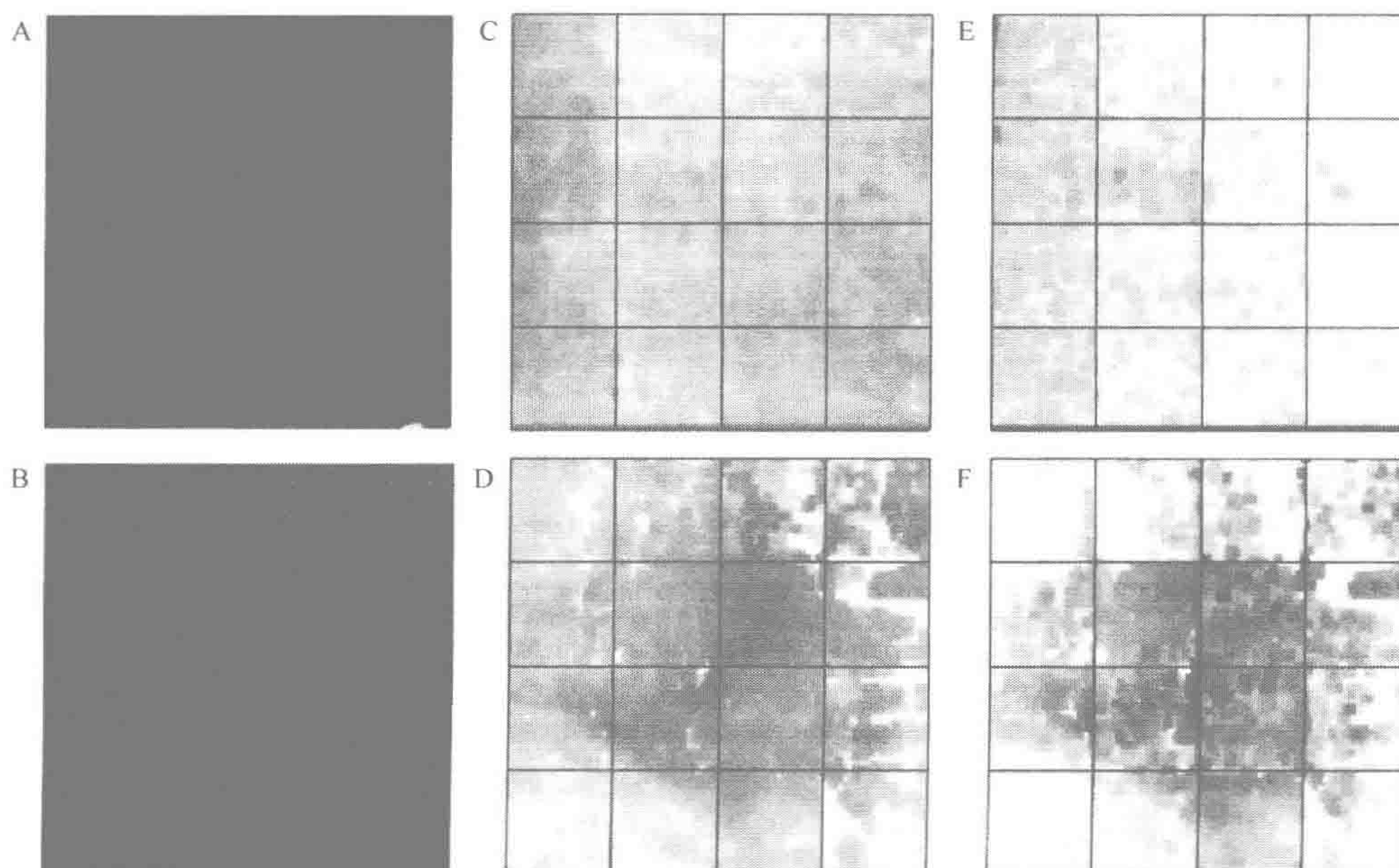
547

图 7-21 盒形图显示为研究突变鼠表型的大量重复抽样试验的对数比率分布
(数据来自 Callow et al. 2000) 图形显示，不同试验的对数比率有不同的中心点，其标度也略有不同 (IQR＝盒尺寸) (见彩图)

空间图

描述背景

整个芯片上的背景通常会发生系统性地变化。使用彩色图像表示测量的背景值为更



548

图 7-22 来自两个不同微阵列背景的空间图
A，B. 两个微阵列的覆盖图；C～F. 高背景区域为伪彩色表示。白色区域代表低背景，分别用亮绿色 (C，D) 或红色 (E，F) 区域代表 Cy3 和 Cy5 通道的高背景 (见彩图)

好地显现背景的变化以及鉴别可能的假象提供了一种简单的方法。图 7-22 中显示了一幅红-绿覆盖图和相应的对各个通道背景的表示。

使极端比值突出

一个芯片上的对数比值出现突出的空间效应并不罕见。对此进行检查的一种简单方法是得到一个关于芯片上样点的示意性的空间表示，并且对分别居于 M 值的上部或者下部 5%的样点使用彩色编码（图 7-23）。信号的非随机分布可能是因为微阵列上那些倾向于高度表达的探针的系统性有序排列。例如，阳性对照通常放置在每个点样针组（pin group）的最上面一行或者数行，并且在这种情况下，它们在空间图上会突出表现为不寻常的亮点。另一种情况是，信号的非随机分布可能是因为杂交假象（见 Part Ⅲ）。观察到的假象的典型图式包括整个芯片上清楚的梯度或者高或低对数比值聚集的区域，通常是沿着芯片的边缘。

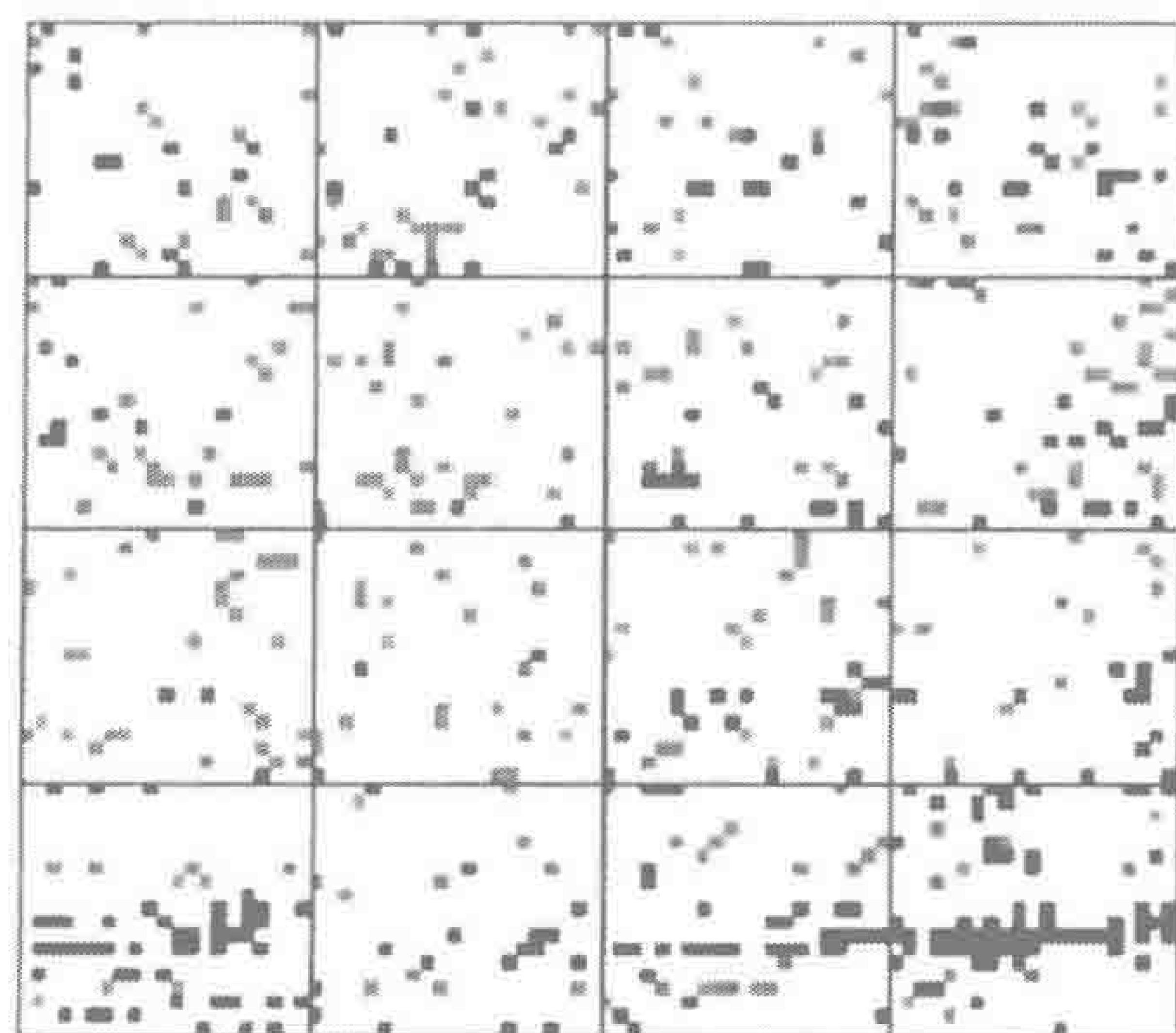


图 7-23　用空间图来表示极端值，如 $\log_2$ 比值的上部（黑色）和下部（绿色）5%部分。在靠近右下方有一显示了高强度信号的非随机分布（见彩图）

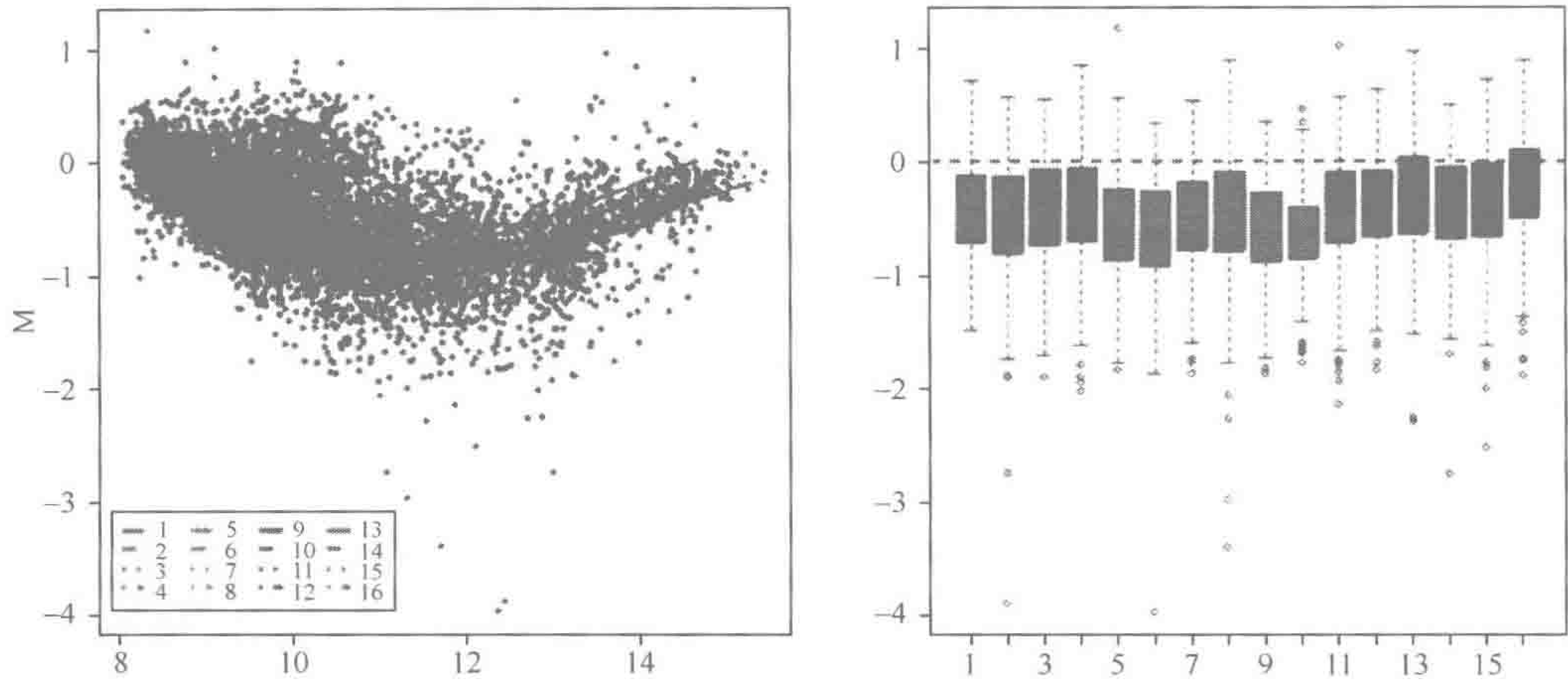
作图工具揭示出的潜在假象

点样针（子阵列）效应

芯片上的样点排列方式对应于用来点样的点样头的形式。例如，点样头由 16 根针排列为 4×4 的阵列形式，并且在一个正方形的阵列中每根针使用了 25×25 次，那么芯片上的样点会以 16 个子阵列，并且每个子阵列包含 625 个样点的形式出现。一个样点所从属的特定子阵列被称为它的点样针组（pin group）。在一个点样针组中的所有点是由点样仪中点样头上的同一根点样针所点样。

样点所从属于哪一个点样针组之中包含了有用的信息，可以是直接的，因为它反映了点样 DNA 时使用的针；或者是间接的，可以用作芯片上样点位置的替代。例如，点样头为 4×4 的点样针阵列时，点样针组 4、8、12 和 16 对应于芯片上的同一边，点样针组 1、2、3 和 4 以及 13、14、15 和 16 同样如此。这在根据点样针组描述数据以及对点样针组内部的偏差进行调整时常常是很有用的。

正如前面所述，描述点样针组效应两种很有价值的作图是 MA 作图和 M 值的箱线图，但是只用于点样针组内部，而不是整个芯片（图 7-24）。针对每个点样针组数据得到的 LOWESS 线（图 7-24，左）显示了红/绿比值当中依赖于强度的变化是不同的，这表明了点样针组偏差。有关 LOWESS 线的进一步说明见 PART Ⅲ。



549

图 7-24　16 条彩色线是通过 16 个点样组所属各组样点的平滑 LOWESS 线（左）和各个点样针组的对数比值的箱线图。这些作图说明了各点样针组之间的差异（右）（见彩图）

加样板效应

cDNA 微阵列的点样通常从在 96 孔微量滴定板上利用聚合酶链反应（PCR）扩增（探针）cDNA 开始，一般每次 4 块板。然后滴定板上的 cDNA 阵列被转移至 384 孔板中，并使用这些 384 孔板逐块进行点样。因此，如果一块芯片中包含由一个 24×24 样点的点样针组构成的 12×4 阵列，那么在每个点样针组之中，每块 384 孔板都会为其中 384/(12×4)＝8个样点提供 cDNA。此外，每组 8 个样点可以归属回到特定的 96 孔微量滴定板。点样过程顺着 384 孔板进行直至结束。

正如上面在“使极端比值突出”中所介绍的，加样板效应也许是探针在多孔板中系统性地有序排列的结果，或者是由于探针制备时的差异。板与板之间 cDNA 探针的质量会有明显的差别，这取决于 PCR 模板是否为质粒 DNA 或细菌培养物，以及与制备 DNA 时的许多步骤有关的差异。有时候，彼此间来源差别非常大的探针会一起放在同一张玻片上。不管是出于何种原因，不同组加样板之间的样点强度特征可能会有看得见的差异。

尽管设计了竞争性杂交来消除探针数量和质量上变化的影响，有时基于加样板效应的系统性偏差仍会存在。检查加样板效应最简便的方法是使用对数强度或对数比值进行平行的箱线图作图，并放置在相同的版面上（图 7-25）。

点样时间效应

点样（例如，一个包含了 1～200 块玻片的批次）会分数天来完成。在此期间，实

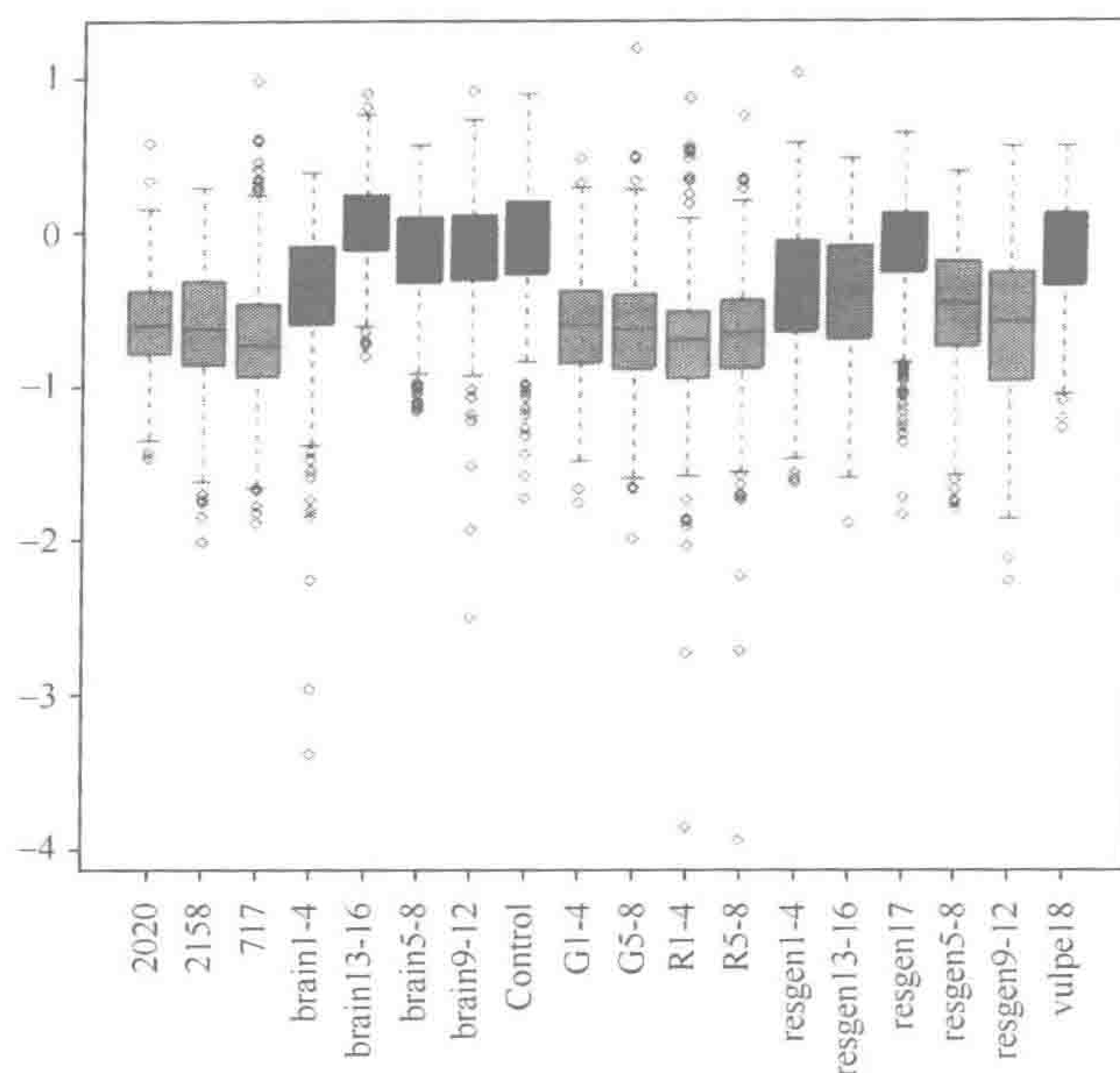


图 7-25 每个箱线图代表了点样在同一块芯片上的 18 个 384 孔板之一的对数比值

从图中可以看出，不同组的加样板之间比值的中值差别很显著。这种类型的差别在进行分析之前应该去除

550

验室中的温度和湿度等条件可能会有较大的变化，灰尘会发生累积。由于这些和其他的因素，不同日期点样的样点有时会在形状、信号强度、样点质量上出现明显的差别。当计算强度比值后，这类变化中的许多会消失掉，但是，保留下来的有时会显而易见（图 7-26）。

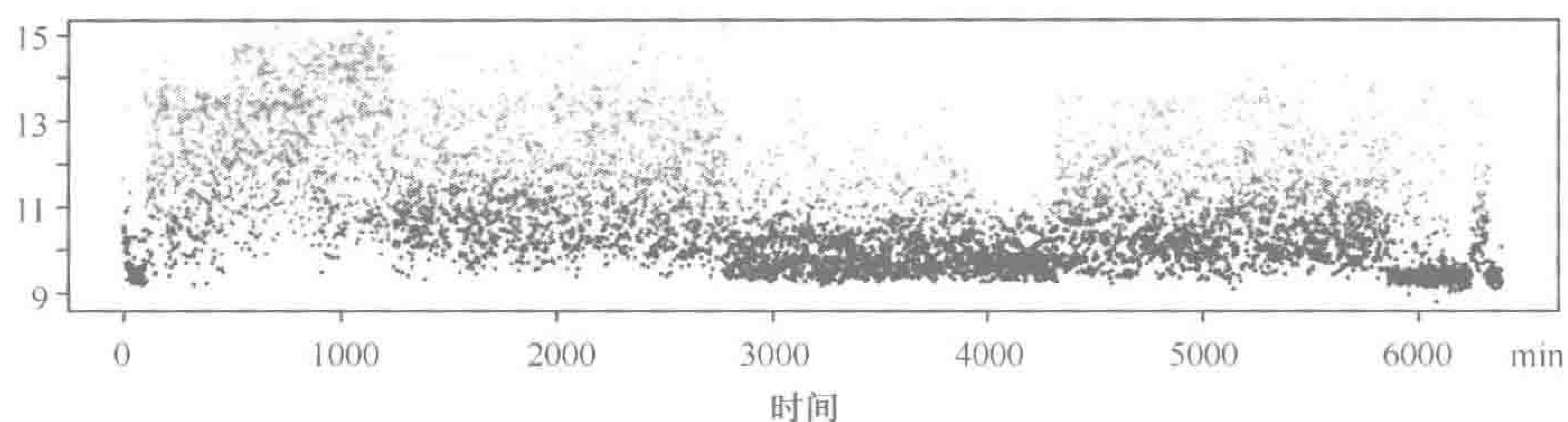


图 7-26 花费 4.5 天完成点样的阵列上样点在绿色通道的强度

x 轴表示样点数目，它在玻片上随着点样时间延长而累积，y 轴为 $\log_2 G$。注意在 4.5 天当中强度的分布是非随机性的。在第三天点样的样点，信号的强度值特别低（见彩图）

551

总结

总之，诸如散点图、箱线图和空间图等探查性的图形工具提供了对微阵列数据快速和简便的察看。这些技术可以表明芯片和实验的质量，因此有助于鉴定可能的系统性偏差。偏差包括了染料间的不平衡、芯片上的空间差异、点样针差异、加样板效应或点样

时间效应。数据的标准化（见 PART Ⅲ）提供了消除这类系统性效应的方法，从而使真正的生物学差异能够更加容易地区别出来。

运用质量测量

许多图像分析程序集中了关于每个样点广泛的定量测量。诸如面积和周长（以像素为单位）一类的形态学度量，以及均一性度量，如样点中像素在每个通道的前景与背景强度和强度比值（进行或者不作背景调整）的标准差。其他的定量度量包括样点亮度指标，如样点前景与背景强度的比值，以及强度超出背景或它的给定倍数的像素在前景中所占的比例。根据这些初始的测量，可以计算进一步的衍生测量，如形状的一种表示（如面积/周长^2）和关于信噪比的其他度量形式，如变异系数（标准差/均值）以及其他的量。Wang 等（2001）根据样点大小、信噪比、局部背景的变异、过高的局部背景和信号饱和等参数提出了一种复合性质量（q_{com}）度量方法。这些作者指出 q_{com} 是反映芯片变化性的一个有效指标。在可公开获取的 Matarray 软件包中实现了这种方法，该软件同时还包括了用于网格化、分割和强度提取的图像分析软件。

微阵列数据的质量评价不仅包括对单个样点性质信息的测量和使用，也包括全局性质量评价的测量。从整个玻片上的样点在两个通道的每一个中对数强度值的分布可以发现这一点（图 7-20）。像素强度通常按比例换算成为在 1 和 $2^{16}=65\ 536$ 之间，也就是说，在 $\log_2$ 尺度上的 0 和 16 之间。观察不到强度利用了满刻度或者几乎是满刻度，通常表明出现了问题，可能会是杂交失败。

我们应该如何利用各种不同的质量度量方法呢？大多数图像分析程序中包括了在一个或多个质量指标的基础上对样点进行标记的方法，常用的指标包含有对信噪比和（或）面积的测量。用户一般忽略掉在最初分析中被标记的样点。类似地，Wang 等（2001）提出 q_{com} 可用来决定是否要设置阈值以剔除那些具有差的复合性质量的样点。正如一系列的可视化检查所表现出来的，这类数据过滤明显改善了数据的外观。但是，当为包含数据而设置阈值时所使用的截止值带有一定主观性时，从统计学的观点来看，把数据作为不满意的而拒绝是困难的。在理想情况下，质量测量中的阈值和权重所选取的数值应该是经过验证能够预测精度的那些质量度量的函数，如芯片间对数比值的方差。除了设定一个明显的边界以包含数据外，更好的方法是以一种渐进的方式给低质量的样点较小的权值。

552

PART Ⅴ　LIMS、数据库和数据管理

Catherine A. Ball，Gavin Sherlock（斯坦福微阵列数据库，斯坦福大学医学院，加利福尼亚，94305-5120）

导言

对基因组进行全面的表达研究这一能力（Lockhart et al. 1996；Schena et al. 1996；Derisi et al. 1997；Lashkari et al. 1997；Wodicka et al. 1997；Spellman et al. 1998）使得许多生物学家对微阵列技术产生了极大的热情，但是会有很多的步骤能让初

次使用者变得畏缩不前。对这类研究产生的数据进行仓储式存储以及与这些数据有关的分析是特别具有挑战性的。对微阵列实验的恰当记录和客观地确定高质量的数据集合的能力会在很大程度上使得对结果的有意义解释和数据假象的检查变得容易。随着数据的增长，对信息的存储和跟踪能力变得必不可少。从DNA点样的玻片微阵列得到的结果的数据管理问题在PART Ⅴ之中解决。然而，许多信息可以很容易地应用到从蛋白质点样微阵列中得到的数据，或者其他的一些平台，如Affymetrix芯片或尼龙滤膜阵列。

微阵列实验生成了大量的数据，这不仅是指扫描芯片输出的数据，而且首先要记录实验是如何进行的：关于微阵列上所使用探针①特性的信息、芯片制备的方法和每个实验的杂交条件。成功地使用DNA微阵列技术要求研究人员能够有效地检索信息和搞清楚这些信息的意思。没有良好的数据管理，几乎不可能查找和解决诸如样品污染和对点样信息的错误记录等问题。此外，除非将基本的信息全面地分类记录，否则其他研究人员就不能对工作进行重复或是评价其质量。尽管许多微阵列的使用者只希望进行有限的一系列实验，但仍然可以从那些正在为大规模的项目进行数据跟踪和存储的用户那里吸取教训。

绝大多数大规模的微阵列研究使用一个LIM（laboratory information management，实验室信息管理）系统来跟踪微阵列是如何制备的，同时使用一个数据库来存储每个阵列杂交结果。结果数据库不仅要保存原始实验数据，还要含有足够的生物学注释以使得对结果的分析成为可能。尽管LIM系统和结果数据库之间的区别似乎是非常自然的，实际上，它们之中的每一个所要求的信息类型是有显著的重叠的，所以在两者之间建立接口使得彼此间的数据能够通过该接口关联起来是非常重要的（图7-27）。由于在LIM系统和结果数据库之间信息存在显著的重叠，因此较小型的项目可以使用兼具了它们功能的单个数据库，这是合理的。 553

从 pubs. acs. org/subscribe/journalsancham/72/supmat/files/ac57A. fil/LIM-Stable. html 可以获得一部分公开的和商品化LIM系统的清单。

LIM系统应该获取的信息

跟踪描述微阵列是如何制备的数据对于将来解释产生的结果以及在实验过程中进行疑难问题的解决都是至关重要的。例如，当微阵列上某条序列的一个拷贝与相同序列的另一个拷贝之间表现不同，那么其中一个样品可能被污染或是错误跟踪。一个典型的微阵列上具有大量的样点，因此必须用自动的方式来进行检测。相反，某个给定序列的表达谱可能会与一个无关序列有密切的相似性。探针跟踪使得研究人员可以检测到一块微量滴定板上的相邻板孔中序列之间显示出相似的表达谱这一情形。如果没有跟踪信息，这类错误的一致性表达谱将会被解释为具有生物学意义。能够检测可能的污染和其他程序上问题的能力使得从可疑样本生成的微阵列样点可以被标记为不可靠的。一个典型的微阵列上具有大量的样点，因此必须用自动的方式使得这类情形容易被检测到。 554

---

① 作者更愿意使用“报告分子”（reporter）这个术语来指点样在微阵列上的核酸，而且这一术语也应用在MIAME方案中（参见关于在研究人员之间共享为阵列数据的信息栏）。为了与手册中的其他章节保持一致，“探针（probe）”一词仍保留在这里（参见第1章导言中关于探针和靶的命名的信息栏）。

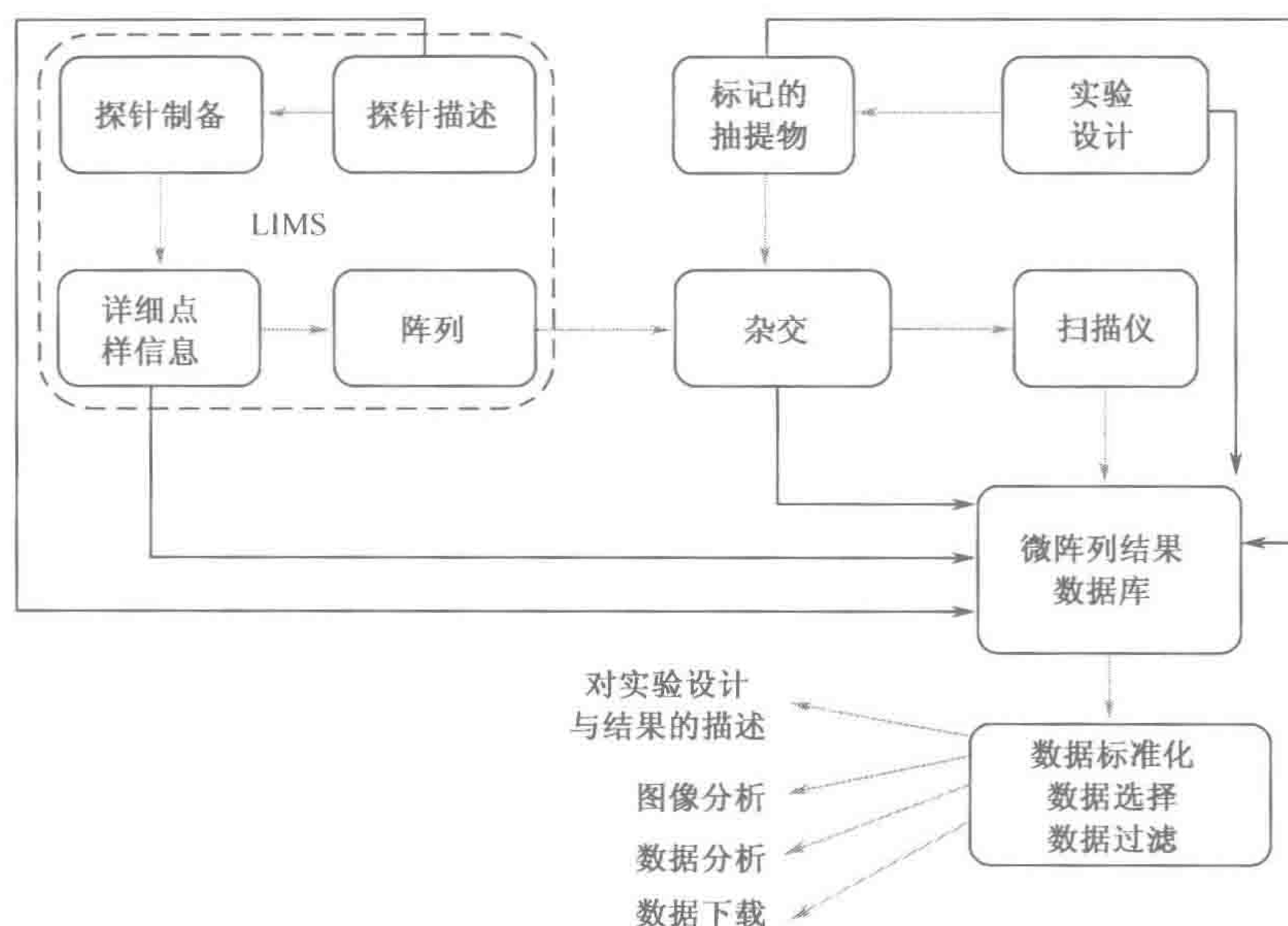


图 7-27 微阵列制备和实验的数据流程图

黑色箭头表明了在微阵列的制备和微阵列实验的设计与完成过程中信息的流动。信息通常存储在 LIM 系统中，包括了对点样在微阵列上的 DNA 分子的描述和制备方法的详细信息，将那些分子点样在微阵列上的信息，以及制备的每块阵列的结局。微阵列结果数据库应该包含以下信息：实验的设计、对探针的描述和杂交中使用的条件以及对杂交后微阵列扫描的结果。灰色箭头表示实验过程的每一步骤中应该保存到微阵列结果数据库以便进行结果解释和疑难问题解答的重要信息

一个有效的 LIM 系统必须能够跟踪产生用于在微阵列上点样的序列的方法，当样品从一块微量滴定板转移到另一块上时要能够映射样品，还要能记录点样在微阵列上的序列的特性以及它们在微阵列上的位置，最后还要能够存储微阵列点样过程的详细情况。

样品的跟踪和制备

对探针的序列特性和它们的制备方法信息进行存储是非常重要的。大多数探针通常是 PCR 产物或者寡核苷酸。PCR 产物可以从许多模板产生，如插入载体的基因组 DNA 或 cDNA。一个微阵列 LIM 系统必须能够记录样品制备的每一步骤以及模板和引物的来源和它们在相关微孔板中的位置。例如，从载体插入片段制备 PCR 产物要求 LIM 系统记录如下信息：

- 克隆/细菌来源（商业性的或者是实验室自己培养的）
- 包含质粒的菌株
- 质粒的特性
- 分离出的质粒的存储位置
- 任何相关的存取号、克隆标识符或者其他从属于质粒中插入片段的标识符。
- 每个引物的性质和来源（通常使用一对通用的载体引物来扩增）
- 每个引物的存储位置

- 每个引物的序列
- PCR 产物的预计大小
- PCR 的性质或者生物学名称
- PCR 产物的存储位置
- 用于评价 PCR 产物质量的任何凝胶图像
- PCR 产物的质量，关于它是否接近预计的大小和（或）是否看见单一或者多重条带，或者根本没有可见的条带。
- 样品在多孔板之间的转移，包括多孔板之间的映射、用于转移样品的技术以及进行转移的日期。
- 制备微阵列的每一步骤中所使用的方案。

能够重建用于生成微阵列中样点的方法对于疑难问题解答是有利的。只有对全部的原始样品检测后才能证实可疑的污染。

阵列上的序列

LIM 系统一个显而易见的用途是明确在阵列上的哪个位置点样了什么。其中一个组成要素是记录设想的探针是什么。对于点样在微阵列上的寡核苷酸，最简单地就是记录它的序列。

但是，如果要点样 PCR 产物，它的唯一性应该基于模板和引物，两个组成要素都必须被正确地跟踪，作为探针性质的一部分。探针在微阵列上放置的坐标也必须和它的性质相随在一起。

微阵列点样

555

微阵列制备在程序上的细节应该包括使用的点样仪的型号（如果可用的点样仪多于一台）、使用的点样针的结构类型、点样针每次加样之间清洗和干燥的次数。大多数研究项目中记录了所制备微阵列的数量、同一批次内每块芯片的顺序（通常在一个批次开始时点样的玻片微阵列上的样点比后来点样的微阵列上的样点要大）、使用的玻片或者点样表面的类型，以及在开始点样之前对表面进行的任何预处理。此外，日期、时间和负责点样的个人的名字常常是有用的。大型设备也将跟踪每块微阵列的结局（是否送给其他研究人员或者只是被放弃和破碎了）以及记账和付款记录。

对结果数据库的要求

当一个微阵列数据库被选定或者研制时，必须保证下列每个问题都得到充分解决。

数据存储

数据必须能够被安全地存储和容易地装入。在选择结果数据库时，重要的是要考虑到并非所有可以利用的数据库都能够合适地接受来自全部微阵列平台的数据。一些数据库软件包能够有效地存储来自玻片微阵列的数据，但是却不能保存来自 Affymetrix 寡核苷酸阵列或者尼龙滤膜阵列的数据，因为这些技术分别提供了不同的数据类型以及不同类型和数量的相关联图像。此外，选择保存微阵列数据的数据库必须和图像分析相匹

配。可用于从扫描的图像中提取数据的免费或者商业软件包有 Scanalyze、GenePix 和 ImageGene。这些软件包生成不同的数据格式以及不同类型和质量的数据点。在理想情况下，来自不同软件包的数据在结果数据库中应该互相一致，这样当不同的软件包提供同类型测量时，可以将结果记录下来。

在结果数据库中与数据存储也有关联的另一个问题是数量庞大的数据。考虑到一个包含 20 000 个样点的微阵列会产生大约一百万个离散的数据点，很显然即使是处理来自一个中等规模实验系列的数据也不是一件普通的任务。大多数图像分析程序为每个样点生成许多测量值，如每个通道的强度、每个通道的背景以及回归相关系数等。相对表达水平可以简单地表示为扣除背景后经过标准化的信号强度的比值。然而，其他的数值（如回归相关系数）可以用作决定数据是否可靠的过滤标准，因此这些数据对于后续的数据分析也是非常关键的。保存关于每个样点的全部测量数据就可以在稍晚一些时候对数据的质量进行评价，或者是让用户可以应用那些即将被开发出来的分析工具（例如，对芯片数据进行标准化的新方法）。这些问题也都与在公共数据库中存储微阵列数据有关，并会在关于研究人员之间共享微阵列数据的信息栏中进行深入的讨论。

556

数据检索和过滤

数据检索

由于数据库能够快速和有效地查询和检索数据才使得它是如此强有力的。一个微阵列数据库必须能够为查询提供一个独立的文件，其中包含了保存的关于微阵列的所有信息，以及与微阵列上探针相关的生物学信息。为阵列上的各个探针进行注释对于结果的解释是非常必要的。不仅如此，一个微阵列数据库必须能够从多个阵列中检索数据，这样来自在实验上有关联阵列（例如，分别是时间序列实验一部分的阵列，或者对相关组织样品间的研究）的结果之间可以作为一个数据集合来进行比较和分析。

数据过滤

在进行数据检索时，应该能够将来自质量可疑的样点的数据过滤掉。对数据质量的度量由实验者来决定，并且也取决于数据库内存储的数据。此外，来自于微阵列制造者的质量度量（PCR 产物是否表现出异常或者样品是否可能被污染）应该从跟踪阵列合成的 LIM 系统传递到所选择的结果数据库。关于数据库进行过滤时所希望的衡量标准的例子有信号对背景的强度、回归相关系数以及标记状态（经由用户或者软件识别的某个样点是不可靠的一种标志）。

一个结果数据库应该具备基于逐个基因进行数据过滤的能力。例如，一个研究人员可能想要在一组阵列当中检索关于一个基因集合的全部数据，或者是在一个阵列集合中选择出那些仅仅与基因发生给定量变化有关联的数据。另一个选项是根据为一个给定的基因提供“良好”数据的微阵列所占百分比来进行数据过滤。一个强有力的结果数据库应该提供多种过滤排列。

生物学样品的数据建模

为了对微阵列实验的结果进行解释，必须将点样在微阵列上的序列以具有生物学意义的方式表示出来。一个简单的模型是假定点样在阵列上的每个序列是一个基因并对它

这样存储。尽管这是表示序列的一个简便方法，但是这种方法具有明显的缺点。例如，这种方法无法让数据库来表示非编码序列，如基因间区或者内含子。此外，一个一对一的、序列到基因的数据模型将不能允许微阵列中包含有表示同一基因的两个截然不同的探针。因此需要一个更为复杂的模型以具备更大的灵活性和对数据质量进行分析。

对数据一种更为复杂的表示方法要求每个序列都拥有自己的标识符。如果一个基因在微阵列中由一条以上的序列来表示，那么每条序列都应该映射到该基因上。这种类型的数据库应该允许来自映射到相同基因的序列的结果可以分别处理，或者是在数据选择时紧缩为一个复合值。

557

使用多个序列来表示一个单独的基因使得研究人员能够发现潜在的假象。一种更加谨慎的策略是不仅跟踪微阵列上的每一条序列，而且还要跟踪该序列的每一个事件。例如，使用 LIM 系统提供的信息，每个 PCR 产物都可以指定一个唯一的标识符。拥有这一信息以后，在某一次生成该序列的实验中加样孔的污染或者是 PCR 的失败都不会导致该序列的其他与本次实验无关的事件被认为是“错误”序列。因此，在选择或者设计数据库模型时，用于追踪微阵列制造信息的 LIM 系统和存储杂交后微阵列结果的数据库之间的信息传递将是重要的考虑事项。

生物学注释

微阵列研究的最终目标是要能够得到关于生物学系统的结论。如果不将微阵列上的序列与基因联系起来，再将基因和科学文献中当前的注释联系起来，这将是不可能的。为了使其价值最大，生物学注释应该是准确的、反映了当前对基因产物最新的理解，而且要与当前的研究有关，并且是简洁的。当研究数百个，甚至是上千个的基因时，弄清楚冗长描述的意思或者借助需要链接到外部的网站都是非常不切实际的。因此希望将基因的注释存储在数据库之中。

理解生物学注释应当是动态的这一点非常重要。通过实验发现或者对基因组的更好理解都会带来新的信息。例如，NCBI 的 UniGene 项目将 EST 聚集到许多聚类的类别。新的序列数据将导致 UniGene 聚类的新版本，并且可能导致版本之间 EST 克隆被关联到不同的基因聚类。为了使得微阵列数据库中的序列注释成为动态的，非常希望数据库软件包具有可以让研究人员经常更新生物学信息的性能。与微阵列上基因的注解相关的问题将在关于基因注释的信息栏中讨论。

实验信息的记录：MIAME

一群微阵列研究人员已经发展了一套标准用来系统地定义，使得其他研究人员可以进行重复实验或者正确地重新分析已发表数据所需要的最少量信息（Brazma et al. 2001）。正如在研究人员之间共享微阵列数据的信息栏中所介绍的，MIAME（minimal information about microarray experiments，微阵列实验的最小限度信息）标准是用来促进数据的共享。可以对公开的数据以一种新的方法进行分析，从而对正在进行的研究起到帮助。相反，其他研究人员对某个研究人员的数据进行创造性地分析也将是有益的。

558 **数据必须能够以软件分析程序所使用格式进行检索**

与分析工具接口

分析工具与结果数据库直接连接常常是方便的或者甚至是必需的。数据分析工具，如层次聚类、自组织图和主成分分析可能会作为数据库软件包的一部分而提供。当然，在使用各种独立的商业性和免费工具之一时，一个显而易见的要求是数据库要能够产生具有正确格式的数据，以便与想使用的工具接口。

LIM 系统和结果数据库之间的接口

在对微阵列数据进行检索和过滤时，为了使得 LIM 系统中的信息有用，在两个数据库之间应该有一定数量的冗余存储。很明显，冗余越多，当数据需要更新时问题就会变得越严重，因此只有那些对微阵列结果进行匆忙过滤和分析所需要的数据才应当被复制。一般的，这主要是在微阵列点样中所使用加样板的固有信息，和这些加样板中包含的样品信息以及由 LIM 系统生成的一个独特的标识符，用于跟踪每个 DNA 片段的每一个事件。两个系统之间的接口可以通过 LIM 系统向结果数据库输送一个点样文件来完成。随着实验输入到结果数据库，它们就可以和正确的微阵列点样文件关联起来。

与结果数据库相关的计算问题

微阵列数据库的存储能力是其中可以容纳实验的数量、每个阵列的大小以及为每个微阵列所保存信息的深度的函数。特别地，如果要将扫描阵列获得的原始数据 Tiff 图像存储起来时，需要的存储会大大地增加。例如，在 2002 年 3 月，斯坦福微阵列数据库（SMD）包括了超过 22 250 个实验，这些结果的记录占用了约 440 000 000 行（每个实验平均 19 750 行）。从物理开销来看，这些数据占用了约 100 GB 磁盘空间，另外有 50 GB 用作结果表的索引。存储扫描每个微阵列生成的原始 Tiff 图像需要约 40 MB，总计将近 1 TB 的存储量。在本书写作时，SMD 中的图像冗余地存档到磁带之前，在 RAID 磁盘阵列上存放的最长时间为三个月。尽管许多研究项目并不会每年产生数千的微阵列，但是研究人员必须预料到将来的需求，并且为各种规模的项目做好相应的计划。

维护一个大型的微阵列数据库可能是一项昂贵的任务。在开支方面要考虑的主要因素有：

- 用于安装和容纳数据库的计算机设备。
- 数据库后端软件。
- 维护数据库的人员。

项目的规模和优先选择的数据库生产商是选择容纳数据库的计算机设备时主要的考虑事项。Unix 通常比其他平台提供了更好的稳定性，并且可以是从运行 Solaris 的 SUN 服务器到便宜的运行 Linux 的 PC，或者运行 MacOSX 的 PowerMac 之中的任何一个。

559 安装微阵列数据库的计算机设备类型部分取决于数据库并发用户的总数。如果数据库是一个多用户的系统，并且许多用户会同时进入数据库进行检索和分析数据，那么只要数据库管理系统（DBMS）可以支持，强烈推荐采用安装多处理器的计算机。

选择数据库软件

数据库软件本身同时存在商业性的和免费的。著名的商用数据库管理系统，如 Oracle 和 Sybase，要比免费的数据库管理系统，如 mySQL，拥有更多的功能，并能提供技术支持。商用数据库软件昂贵的价格和相比之下公开性解决方案有限的可利用前端软件是必须要考虑的问题。

免费数据库

商用微阵列数据库提供了详尽的文档。对众多这类微阵列数据库的极佳评述，请见 Gardiner-Garden 和 Littlejohn (2001)。下面介绍的三个数据库可以免费使用，至少对学术机构是这样的，并且分发了完整的源代码和设计，允许企业实验室定制（或者改进）。关于它们一些比较重要的性能或者不足在表 7-7 中分别进行了总结。

表 7-7 免费微阵列数据库软件包之间的比较

软件	优点	缺点
AMAD：microarrays. org/software. html	简便，容易使用 维护成本低	不是一种关系数据库 不容易升级
SMD：genome-www. stanford. edu/microarray	可升级为大容量数据集合 对探针进行自动的注释更新 用于数据解释和质量控制的复杂分析工具。开发很活跃，已经被成功地应用	需要 Oracle RDBMS 安装和维护很费力
GENEX：www. ncgr. org/genex	不需要商业化的 RDBMS。可接受来自多个平台的数据。具有数据分析和解释的良好工具 开发很活跃	没有被广泛地应用并且实验相对较少 不能查看被扫描阵列的代理图像(proxy image)

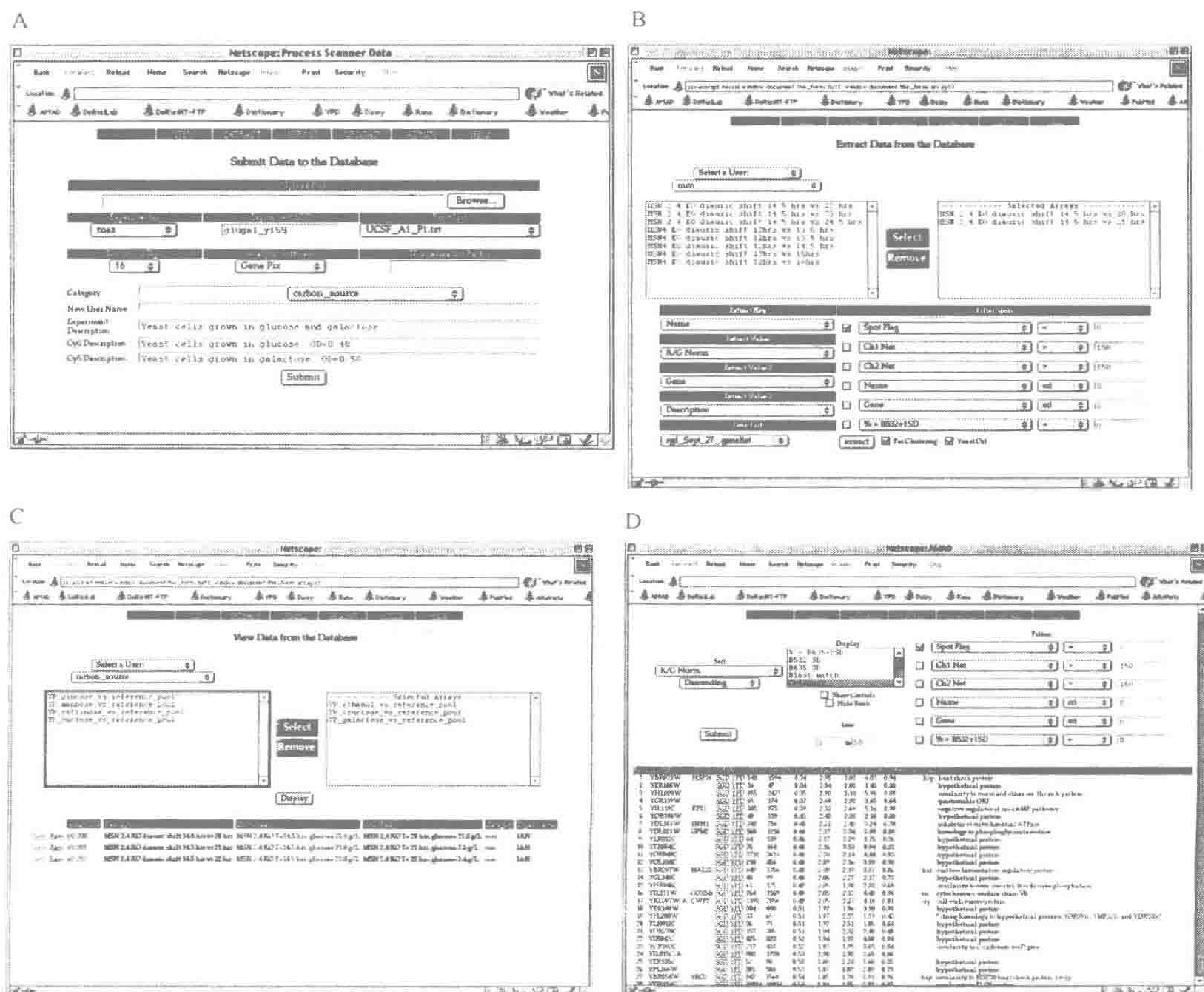
AMAD

AMAD (another microarray database) 是从斯坦福大学 Brown 和 Botstein 实验室内部使用的数据库发展而来的一个普通文本文件数据库。作者是 Mike Eisen，Paul Spellman，Joe DeRisi 和 Max Diehn。AMAD 可以从 http://microarrays. org/software. html 获取。

560

- **需求和安装**：AMAD 要求一台能够运行 Web 服务器的计算机，并且机器中运行了 Perl。一般，正如安装文档声明的那样，一台运行 Linux 的 PC 或者运行 MAcOSX 的苹果机就足够了。AMAD 的安装很轻松，有详细的文档说明如何来完成。
- **性能**：AMAD 允许向文件系统上传普通文本文件。数据可以按照实验进行逐个查看，或者对多个实验进行同时检索并输出单一的文件以用于后续分析。输出文件与 Part Ⅵ 中介绍的聚类程序相匹配。从 AMAD 中检索数据时，文件中的任何一列数据都可以被用来过滤数据。此外，在说明微阵列点样的内容时，用

户可以添加额外的数据列，并且 AMAD 将把这些数据转换为实验数据文件。AMAD 与来自 Scanalyze（http://rana.lbl.gov/EisenSoftware.htm）或者 GenePix（http://axon.com/GN_ GenePixSoftware.html）的数据兼容，这两种软件是分析微阵列扫描图像常用的软件包。图 7-28 中给出了经过选择的一些屏幕截图，说明了 AMAD 的一些数据选择性能。



561

图 7-28 A. AMAD 中的数据提交工具。AMAD 的数据提交工具允许用户从计算机里上传文件，并从列表里选择实验者、点样列表、点样针形状、分析软件，以及实验分类。B. AMAD 中的数据选择。AMAD 允许用户浏览微阵列结果列表、选择子集，并指定检索哪些数据进行进一步的分析。C. 在 AMAD 中察看对微阵列数据的描述。AMAD 的用户可以察看对微阵列实验的描述，并通过工具来察看结果或者只是下载原始数据。D. 在 AMAD 中察看微阵列的结果。AMAD 允许用户对结果过滤，定义数据的显示和对微阵列实验的结果整理排序

- **优点**：对于预算有限，而且数据量在几十到几百个微阵列之间的实验室，AMAD 是一个简单而且有效的数据库。它安装容易，界面整洁，易于使用。
- **缺点**：AMAD 并不是一个关系数据库，因此，也就不具有稳健的关系数据库所提供的方便性能。数据库中有大量的冗余（例如，来自一次点样的每个实验都复制全部相同的点样信息），使得数据更新的开销很大。AMAD 的普通文本文件系统的限制使得该软件包不适用于需要容纳大量的微阵列结果的任何项目。

此外，当点样列表产生错误时，除了重新输入含有正确点样列表的实验之外，没有其他简单的方法可以用来对实验更新。AMAD 已经不再有活跃的开发。

斯坦福微阵列数据库

斯坦福微阵列数据库（stanford microarray database，SMD）（Sherlock et al. 2001）是一个使用 Oracle 作为数据库管理系统的关系数据库。在斯坦福开发的 SMD 是为了取代作为 AMA 原型的最初的普通文本文件系统。SMD 可以从 http://genome-www. standford. edu/microarray 获取。图 7-29 中的路线图说明了 SMD 中的一些数据流和工具。

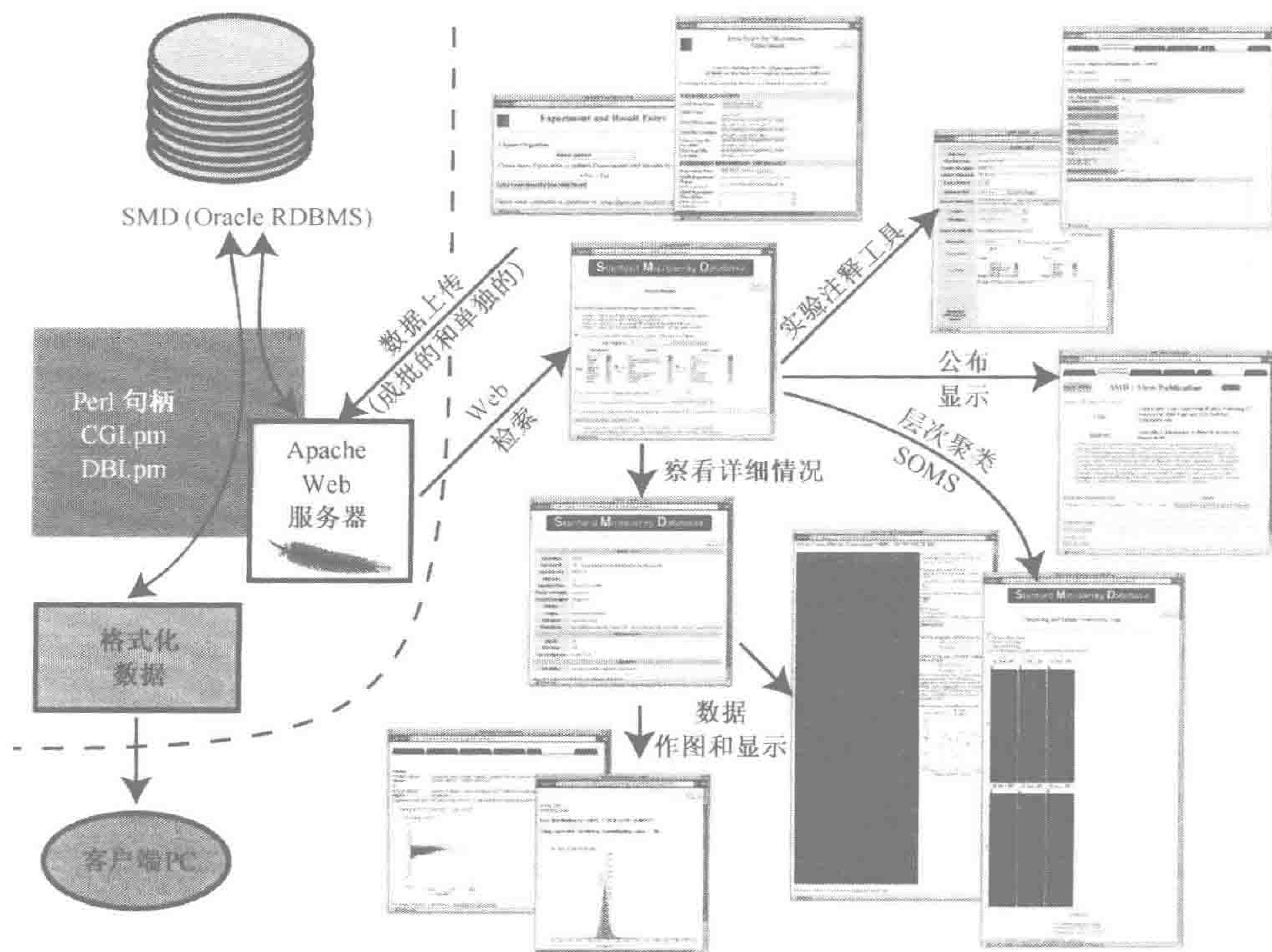


图 7-29　SMD 的路线图

图中说明了数据的流动和用于质量评估与数据作图显示，以及进行下游分析的相关工具

562

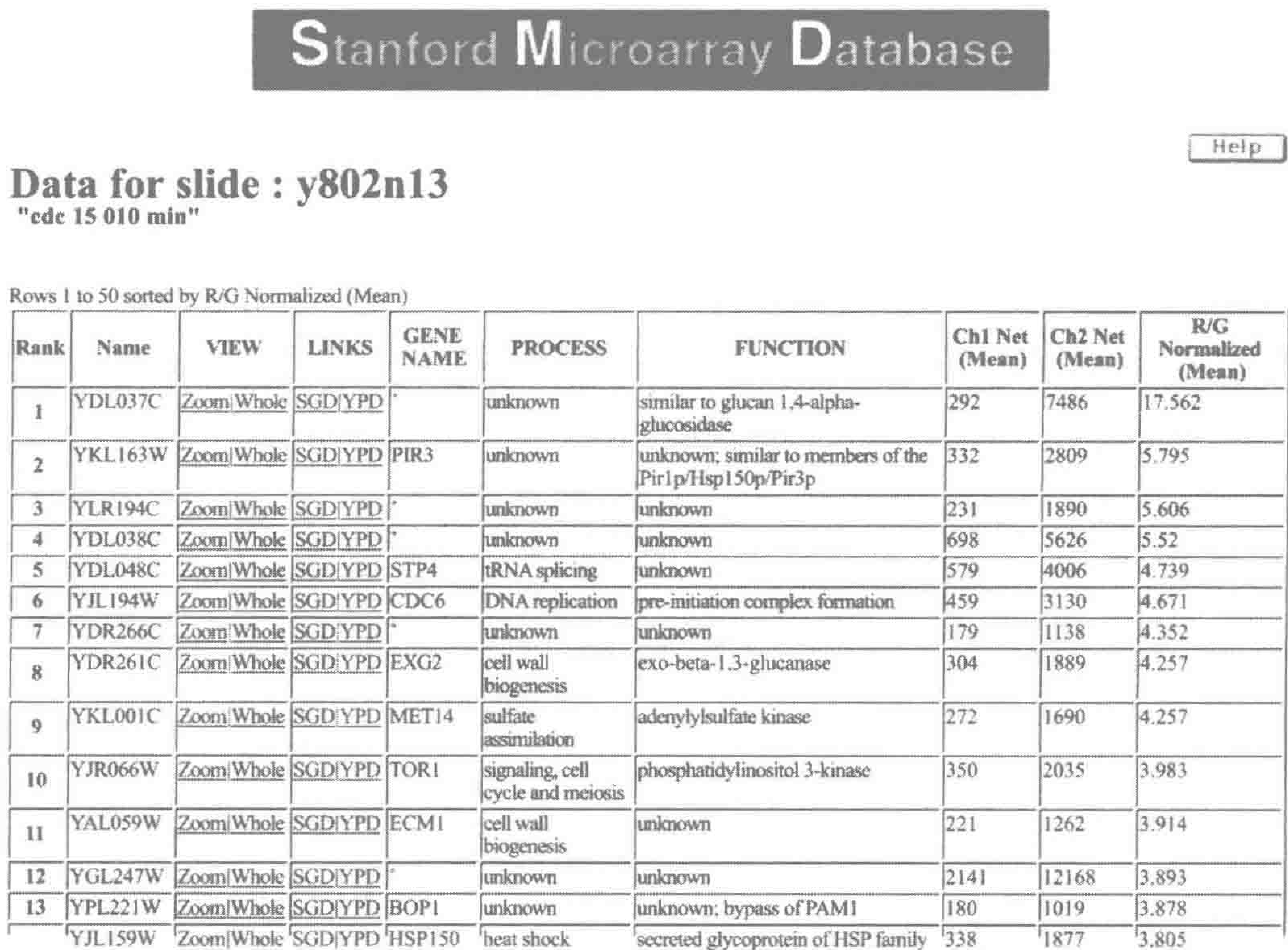
- **需求和安装**：SMD 可以被安装在低端系统上，虽然直到现在，斯坦福的安装是在一台 Sun E4500 服务器上，运行 Solaris 8，带有 8 个处理器和 8 GB 的 RAM。尽管采用 Solaris 系统没有特别的原因，但是注意到在 Oracle 的更新和程序缺陷补丁中，一般是该平台的首先出现。SMD 的安装也要求一份 Oracle 企业版服务器软件的拷贝、一个 Web 服务器、Perl 以及一些 Perl 模块。尽管在随同软件发布的安装脚本解决了使软件运行的大多数细节问题，它的安装当前仍然不是一个简单的过程。而且，要进行完全安装，必须解决许多额外的细节，如建立 Oracle 数据库实例和生成全部表格以及它们之间的联系。

尽管进行这些工作所需要的全部 SQL 脚本已经发布，但是需要一个经过训练的数

据库管理员来使用他们。

- **性能**：SMD 中允许输入来自 GenePix 或 Scanalyze 的数据（以逐个或者批处理的方式），随后可以应用复杂的过滤功能对每个实验或者许多实验进行数据检索，这在很大程度上与 AMAD 类似。另外，SMD 在数据库顶层开发了许多工具，可以用来评价阵列的质量和实验的可重复性，以及微阵列原始数据的可视化。

SMD 还提供了下游分析工具。当前仅限于层次聚类（Eisen et al. 1998）和自组织图（Tamayo et al. 1999，Toronen et al. 1999）。将来发布的软件中会包括 k-均值聚类（Tavazoie et al. 1999）、奇异值分解（Alter et al. 2000）和缺失数据的估算（Troyanskaya et al. 2001）。图 7-29～图 7-34 中，对数据选择、分析、质量控制和生物学注释等部分功能进行了说明。



Rank	Name	VIEW	LINKS	GENE NAME	PROCESS	FUNCTION	Ch1 Net (Mean)	Ch2 Net (Mean)	R/G Normalized (Mean)
1	YDL037C	Zoom|Whole	SGD|YPD	*	unknown	similar to glucan 1,4-alpha-glucosidase	292	7486	17.562
2	YKL163W	Zoom|Whole	SGD|YPD	PIR3	unknown	unknown; similar to members of the Pir1p/Hsp150p/Pir3p	332	2809	5.795
3	YLR194C	Zoom|Whole	SGD|YPD	*	unknown	unknown	231	1890	5.606
4	YDL038C	Zoom|Whole	SGD|YPD	*	unknown	unknown	698	5626	5.52
5	YDL048C	Zoom|Whole	SGD|YPD	STP4	tRNA splicing	unknown	579	4006	4.739
6	YJL194W	Zoom|Whole	SGD|YPD	CDC6	DNA replication	pre-initiation complex formation	459	3130	4.671
7	YDR266C	Zoom|Whole	SGD|YPD	*	unknown	unknown	179	1138	4.352
8	YDR261C	Zoom|Whole	SGD|YPD	EXG2	cell wall biogenesis	exo-beta-1,3-glucanase	304	1889	4.257
9	YKL001C	Zoom|Whole	SGD|YPD	MET14	sulfate assimilation	adenylylsulfate kinase	272	1690	4.257
10	YJR066W	Zoom|Whole	SGD|YPD	TOR1	signaling, cell cycle and meiosis	phosphatidylinositol 3-kinase	350	2035	3.983
11	YAL059W	Zoom|Whole	SGD|YPD	ECM1	cell wall biogenesis	unknown	221	1262	3.914
12	YGL247W	Zoom|Whole	SGD|YPD	*	unknown	unknown	2141	12168	3.893
13	YPL221W	Zoom|Whole	SGD|YPD	BOP1	unknown	unknown; bypass of PAM1	180	1019	3.878
	YJL159W	Zoom|Whole	SGD|YPD	HSP150	heat shock	secreted glycoprotein of HSP family	338	1877	3.805

图 7-30　SMD 中微阵列结果的察看

在 SMD 中可以察看关于每个样点的结果，以及对点样中使用的 DNA 分子的介绍。网页上的链接可以用来查看样点的图像、整个微阵列以及来自外部数据库中的报道

563

- **优点**：SMD 是一个可升级的微阵列数据存储解决方案。在斯坦福安装的数据库中当前容纳了超过 22 250 个实验，包含了来自超过 440 000 000 个样点的数据。另外，SMD 具有灵活的安全模型，允许对数据和工具的使用进行细分的访问控制。数据库本身有很多有用的工具，同时软件允许用户察看微阵列扫描的代理图像，以便对结果质量进行可视化的评估。不仅如此，SMD 对在微阵列上所表示的人、小鼠和酵母基因的注释进行动态更新。SMD 正在进行活跃的开发，因

此将会有新的特性和经过改进的设计与软件。

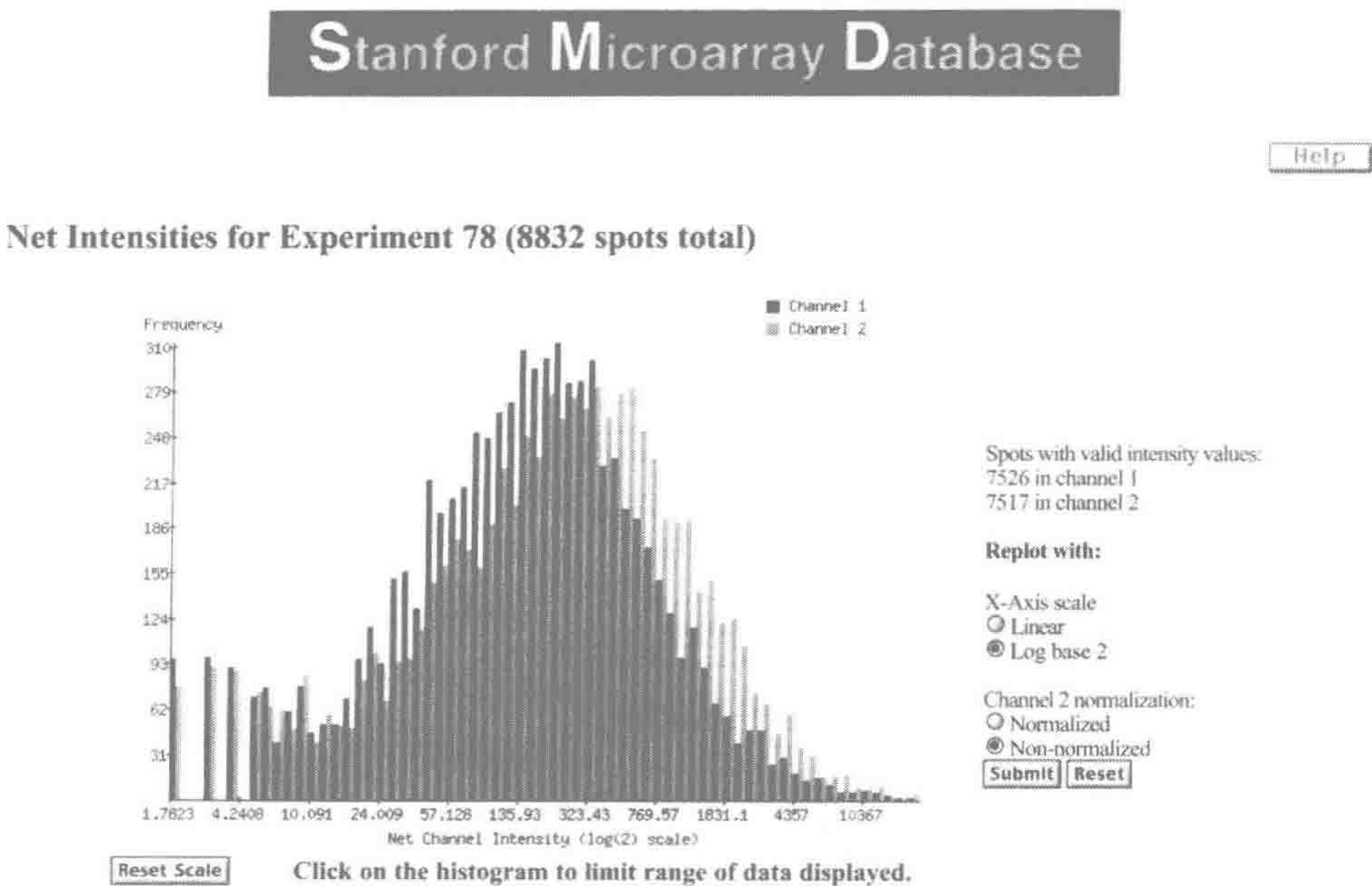


图 7-31　SMD 中的表达变化情况

SMD 的用户可以察看数据库的所有微阵列中独个基因的表达。绘制出强度分布图。用户可以点击直方图中的某个点来查看关于某个基因当强度位于一定范围之内的那些实验

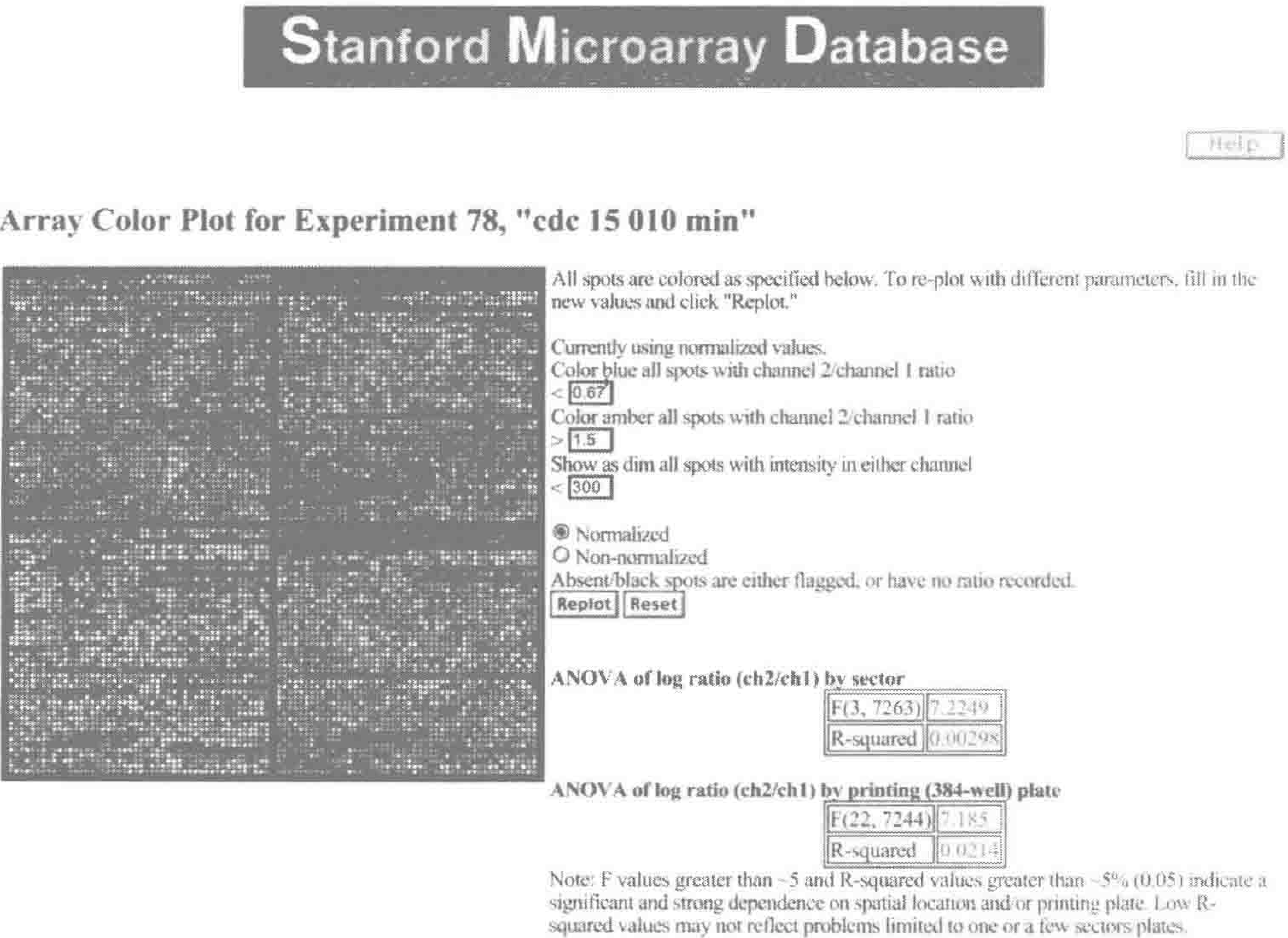


图 7-32　SMD 中阵列的彩色作图

使用显示整个微阵列中样点强度的工具，SMD 用户可以检查他们结果当中存在的空间偏差

564

S.O.U.R.C.E. Search | Switch to GeneReport | Help

	IMAGE:1542757
cDNA Clone ID	IMAGE:1542757
Clone Source	NCI
Organism	Homo sapiens
dbEST Library Number/ Name	1042 [Unigene Lib ID = 600] / Soares_NFL_T_GBC_S1
Insert Length	1024
Vector	pT7T3D-Pac (Pharmacia) with a modified polylinker
	3' EST Information
dbEST ID	1637193
GenBank Accession	AA909121
GenBank GI History	3048526
High Quality Seqence	stops at base:*345
Putative ID	gb:M81590 5-HYDROXYTRYPTAMINE 1B RECEPTOR (HUMAN);
dbEST Entry Created	1998-05-20
dbEST Last Updated	1998-06-17
Perform a NCBI BLASTN	NR ****HTGS ****dbEST
Perform a NCBI BLASTX	NR ****SwissProt
UCSC Genome Viewer	Genome Viewer

| Search | GeneReport | Help |

Feedback for: array@genome.stanford.edu

SOURCE is provided by the Genetics Department, Stanford University.
©2000 Stanford University. All Rights Reserved.
Compiled from a variety of public databases.
Data on this page is curated from UniGene , dbEST , and the Genome Sequencing Center at Washington University, St. Louis

图 7-33　SMD 中克隆的 SOURCE 显示

SMD 为用户提供关于人、小鼠和大鼠的 cDNA 序列编辑好的注释。信息有许多来源，包括 dbEST 和 GenBank

S.O.U.R.C.E. Search | View Clones for this Gene | Help

	TP53
tumor protein p53 (Li-Fraumeni syndrome) UniGene , LocusLink , OMIM , GenAtlas , GeneCard , Ensembl , MapView	
Aliases	
• p53 • PHOSPHOPROTEIN P53 • CELLULAR TUMOR ANTIGEN P53	
Chromosomal Location	
Chromosome/Cytoband	17p13.1
Microarray Gene Expression Data	
Data available	Show Gene Expression Data
LocusLink Information	
Locus Link Summary	Tumor protein p53, a nuclear protein, plays an essential role in the regulation of cell cycle, specifically in the transition from G0 to G1. It is found in very low levels in normal cells, however, in a variety of transformed cell lines, it is expressed in high amounts, and believed to contribute to transformation and malignancy. p53 is a DNA-binding protein containing DNA-binding, oligomerization and transcription activation domains. It is postulated to bind as a tetramer to a p53-binding site and activate expression of downstream genes that inhibit growth and/or invasion, and thus function as a tumor suppressor. Mutants of p53 that frequently occur in a number of different human cancers fail to bind the consensus DNA binding site, and hence cause the loss of tumor suppressor activity. Alterations of the TP53 gene occur not only as somatic mutations in human malignancies, but also as germline mutations in some cancer-prone families with Li-Fraumeni syndrome.
SwissProt Information	
SwissProt Accession No.	P04637 ' CELLULAR TUMOR ANTIGEN P53 (Homo sapiens)

565

图 7-34　SMD 中基因的 SOURCE 显示

SMD 的用户可以从 SOURCE（Stanford Online Universal Resource for Clones and EST）获得有关基因的信息。信息根据许多来源进行编辑而成，如 UniGene、LocusLink 和 OMIM

- **缺点**：SMD 的安装或维护都很困难，它还需要昂贵的硬件和软件以及经过训练的员工（至少有一名数据库管理员和一名程序员/掌管人），以保证其运转。当前，SMD 中还未保存有足够的对实验和样本的注释以满足 MIAME 的要求，虽然这已经被作为将来的目标。此外，对 SMD 所使用设计进行的一些修改需要对基因/位点和微阵列上的核酸序列这一概念，以及它们之间的映射进行更好的建模。SMD 中当前不对单通道的微阵列数据进行存储，包括那些从 Affymetrix 的 GeneChip 或者放射性标记的大型阵列。

GeneX

GeneX 是一个使用免费数据库系统、PostgreSQL 来存储数据的免费数据库（在 GNU 公共许可下发布）。关于 GeneX 的详细信息，请见 http://genex.sourceforge.net/。

- **需求和安装**：GeneX 可以被安装在一台 Linux 计算机上，同时要求其运行 Apache Web 服务器并安装有 Perl。通过使用一个安装脚本来配置系统的各个部分。软件包的总体结构如图 7-35 所示。

567

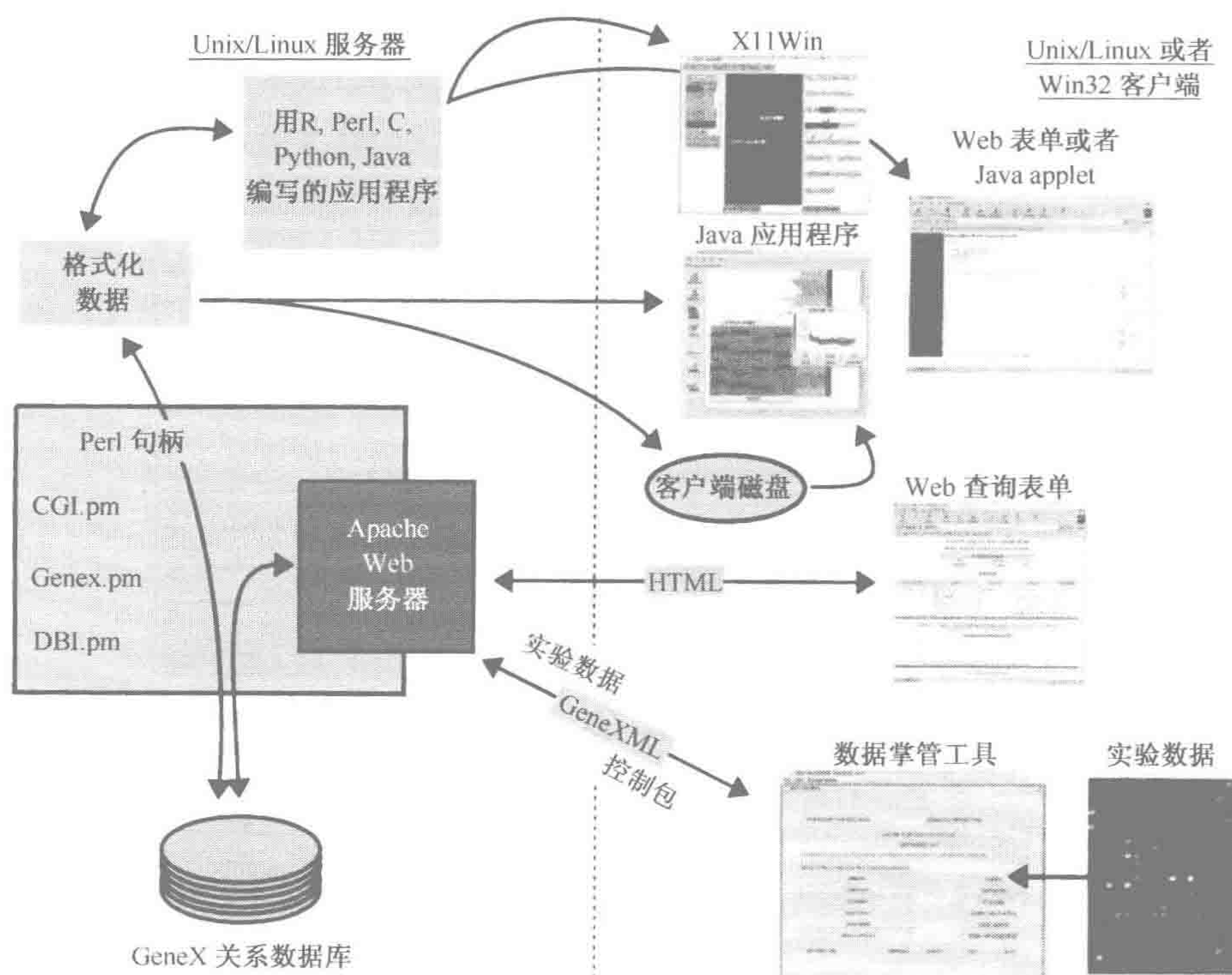


图 7-35　GeneX 的路线图

GeneX 需要一台 Unix 或者 Linux 服务器及 Unix 或 Linux 客户端。数据可通过 Web 界面上载到 GeneX 中。数据库能够向 web 表单，或者是数据库外部的应用程序，如聚类程序输出数据

• **功能**：Genex 提供了一个用 Java 编写的客户端的掌管工具。可以用来格式化数据集合以便于安全地将数据上载到数据库之中；它同时还具有简单的 html 界面，用于在数据集合之间或者是单个实验之中进行数据查询。如果有经验的程序员对界面进行扩展，软件就可以进行更为复杂的查询。对于检索得到的数据，可以运行许多种分析程序，如聚类、多维尺度分析（multidimensional scaling）、聚类有效性验证（cluster validation）和主成分分析。图 7-36～图 7-38 中说明了 GeneX 的一些数据分析功能。

GeneX: xcluster - Hierachical Clustering for Gene Expression Data

General Help   Example Paired Data Set

Data Input Parameters	
Data File to Upload: (Format expected, Data coding) If your file won't upload, or xcluster won't process it, check these possible reasons.	sample.txt [Browse...]
Check this button to take the natural log (ln) of uploaded values * (if your data hasn't already been log-transformed).	☐
File Name Root for Output Files: (only a-z, A-Z, _, 0-9, 28 chars max)	xcluster
Data fields delimited by: * whitespace = TABS & spaces	whitespace*

Data Partitioning. (Optional)		
The BLUE and GREEN sections below are BOTH OPTIONAL and apply to partitioning the data before hierarchical clustering. (Neither will be done unless a button is checked and then only if there are values in the appropriate spaces). The BLUE corresponds to partitioning via KMeans. The GREEN corresponds to partioning via Self Organizing Maps.		
◉	No Partitioning	
○	Number of K-Means clusters to generate	N =
○	Choose an X dimension for the SOM.	X =
*	Choose a Y dimension for the SOM.	Y =
*	Seed the random number generator with the time?.	☑

Variables for Hierarchical Clustering	
What kind of Distance Metric to use?	Pearson Correlat'n
Metric for Clustering based on Genes	Centered *
Metric for Clustering based on Experiments	No Clustering
	* Only applies to Pearson Correlation.

566

图 7-36　GeneX 中的聚类时的选择

通过使用数据库给出的文件，GeneX 的用户可以为层次聚类，包括自组织图和 k-均值聚类进行多种选择。结果可以按照基因和（或）实验进行分组

• **优点**：与 AMAD 类似，GeneX 的安装相当地简便，并且不需要安装或维护昂贵的硬件和软件。它的软件是完全开放源代码的，并使用了其他的开放源代码工

具。它有考虑周到的数据模型，可以存储来自不同微阵列平台的数据，如双色的微阵列数据和单通道的 Affymetrix 微阵列数据。

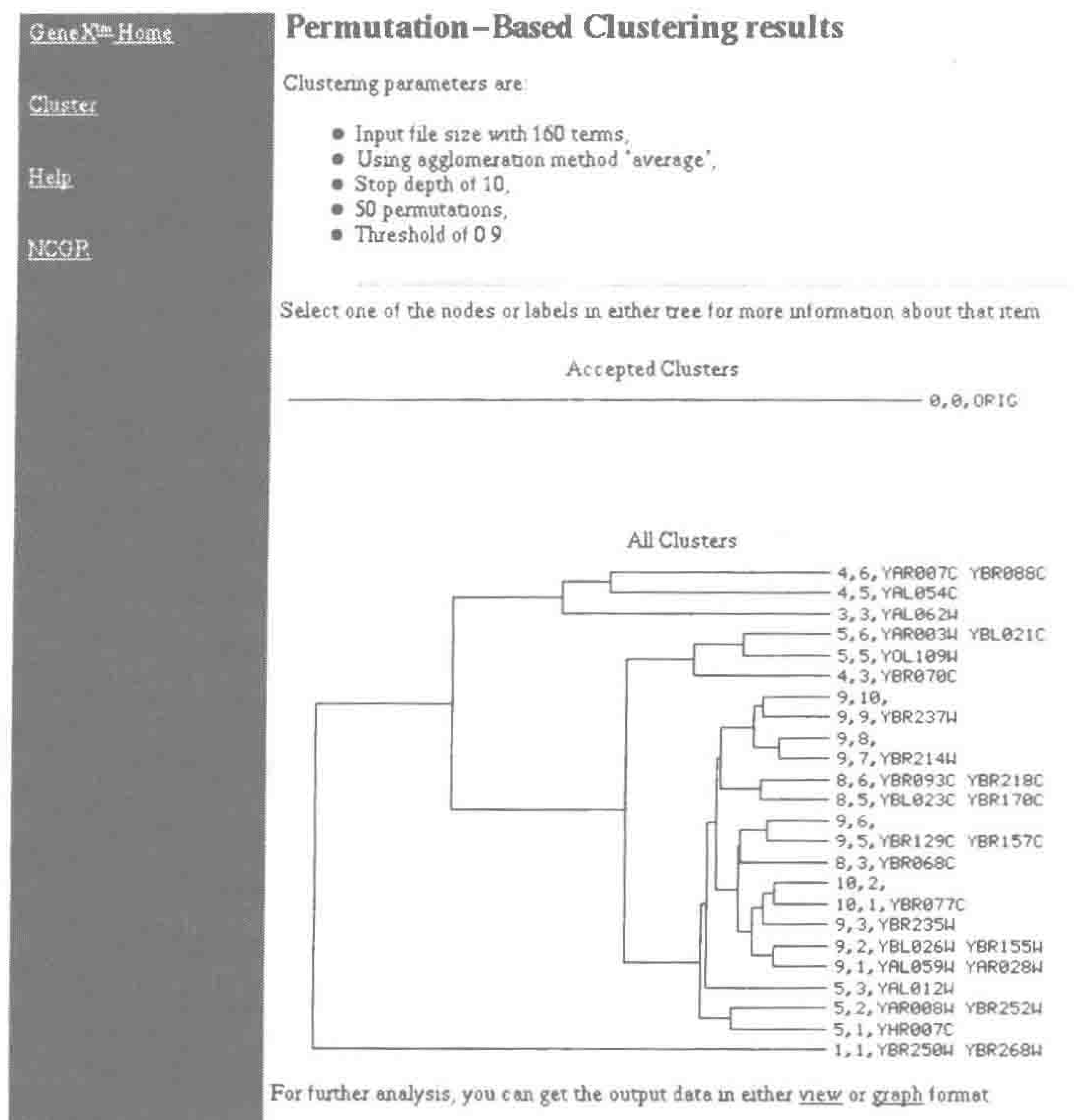


图 7-37　GeneX 中的 r-Cluster

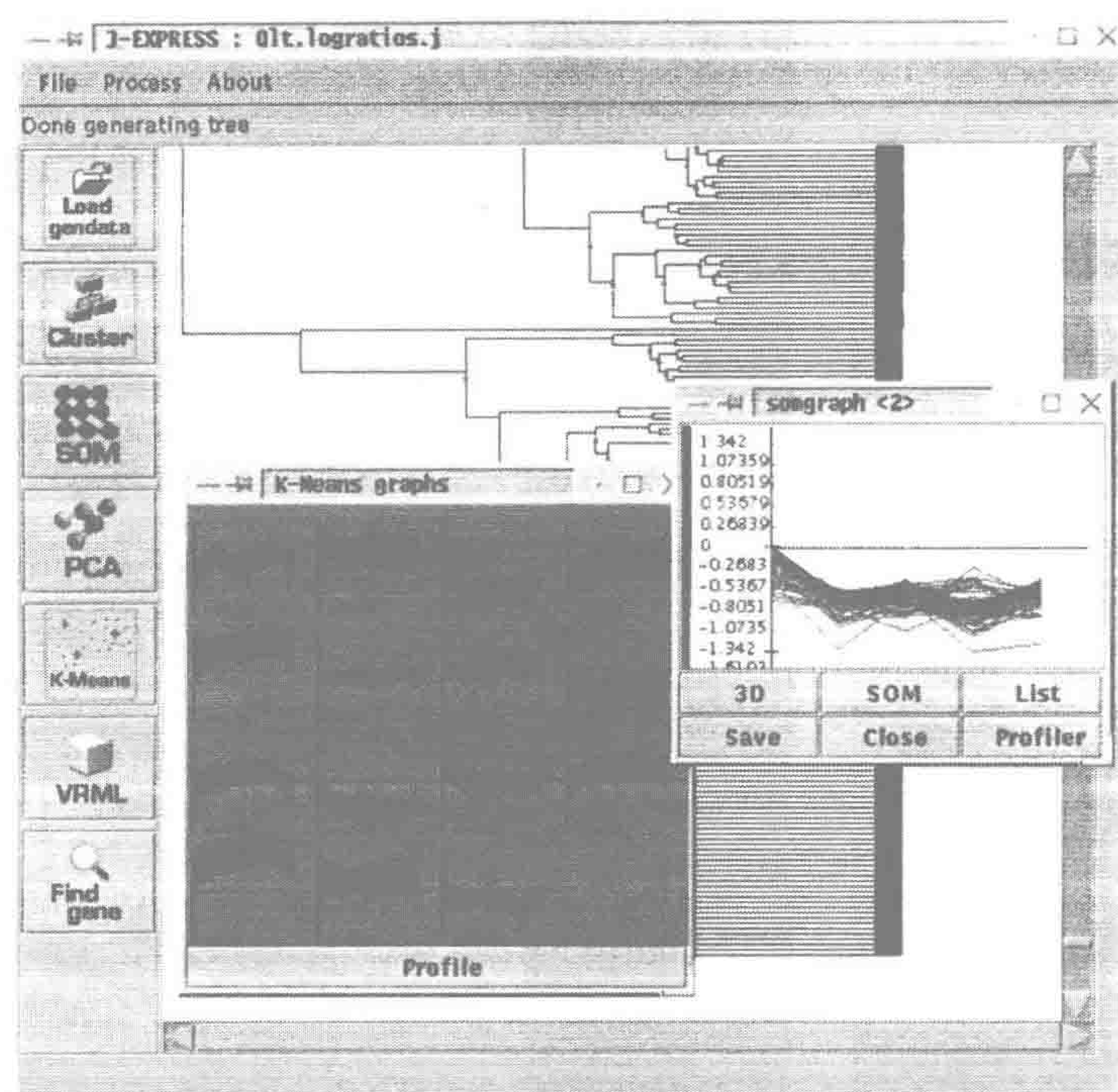
GeneX 输出的数据可以提交给许多分析工具使用，如 R-cluster，一种对聚类后的微阵列实验结果进行显示和浏览的工具

GeneX 提供了大量的数据分析工具，同时 GeneX 处在活跃的开发当中，对软件、数据模型和分析工具都将有不断的改进。

- **缺点**：GeneX 目前还没有被证明可以升级到容纳数千个实验，当然，这并不意味着它不能被升级。当前，GeneX 中不对 Tiff 图像进行存储，并且也不能查看表示微阵列实验质量的代理图像。

致谢

非常感谢 John C. Matese 和 Harry Mangalam，他们提供了一些图片和有帮助的讨论。同时要感谢 Tina Hernandez-Boussard 和 Janos Demeter 对手稿的仔细阅读。



568

图 7-38 GeneX 的 J-Express

GeneX 用户可以使用 J-Express 查看聚类数据的节点上的表达谱图。J-Express 提供了许多工具来帮助实验者浏览他们的聚类数据，如按基因名称搜索和生成 SOM 或者 k-均值聚类

569

PART Ⅵ 聚类分析与显示

Paul T. Spellman（加州大学，伯克利，加利福尼亚，94720-3200）

导言

微阵列的制备和数字信息的产生仅仅是理解通过基因表达谱而获得的基因组水平数据的第一步。简单的比较——例如，识别在一个实验样本中的表达比另一个实验样本中高的基因——可以很容易地通过诸如电子制表软件一类的工具来进行分析。然而不幸的是，大多数的数据集合都不容易研究，因此需要更有力的分析技术。Part Ⅵ 着重介绍两种免费工具的使用：Cluster，用于将相关的基因表达数据组织起来；Tree View，用于使得聚类后微阵列数据的显示容易进行（Eisen et al. 1998）。这里提供的许多信息与其他的微阵列信息分析模式也具有相关性。

聚类工具虽然并不能够完全适用于所有的分析，但它是用来组织基于微阵列的基因表达数据的一种简便而有力的方法。在某些情况下，对结果事先的期望或者实验的目标可能会提示使用其他的分析方法。通常，如果一个实验设计为探索性的，而不是用来回

答某一特定的问题，那么聚类方法将是进行最初数据分析一个很好的选择。探索性研究的例子包括确认在转录水平上受到某些事件，如生长因子刺激（Iyer et al. 1999）调控的基因，或者寻找生物学处理之间的共性，如比较环境胁迫（Gasch et al. 2000）或者是基因分型（Hughes et al. 2000）。当实验的目的是发现与特定疾病状态或临床结果相关的基因时，采用更为关注的（有监督）方法就比较合适了。

有监督和无监督方法

有监督方法通常借助于已经存在的信息来帮助指引数据间的联系。与此相反，无监督方法利用基因表达信息来揭示数据中的模式。在有监督学习中，一组样本先被确定为几个亚型，再借助于该信息来定义区分亚型的基因集合。当定义完成之后，在亚型没有被算法知晓的情况下，测试判别集（discriminating set）对样品的分类能力。例如，研究人员可能会希望使用基因表达数据将白血病样本归类到已知的亚型中去。

无监督方法，如无监督层次聚类，在用于研究数据集以发现事先未预料到的知识，或者在关于样品的额外信息无法获得时，是非常有用的。其他常见的无监督方法包括 k-均值聚类、自组织图和主成分分析。尽管无监督方法看起来要比有监督方法的偏差更低，但是在全部可用信息都得到使用的情况下，识别有用分子类别的能力可能会得到增强。为了使结论更加稳健和可靠，训练集和测试集需要满足很大程度上的独立性。

聚类的目的是为了组织微阵列数据以便更好地认识和研究其中潜在的结构。许多研究论文都表明基于基因表达数据的聚类结果中每个类别都富含了已知参与相似生物学过程的基因，这暗示了未知功能的基因也可能参与了那些同样的过程（Eisen et al. 1998; Tavazoie et al. 1999）。

570

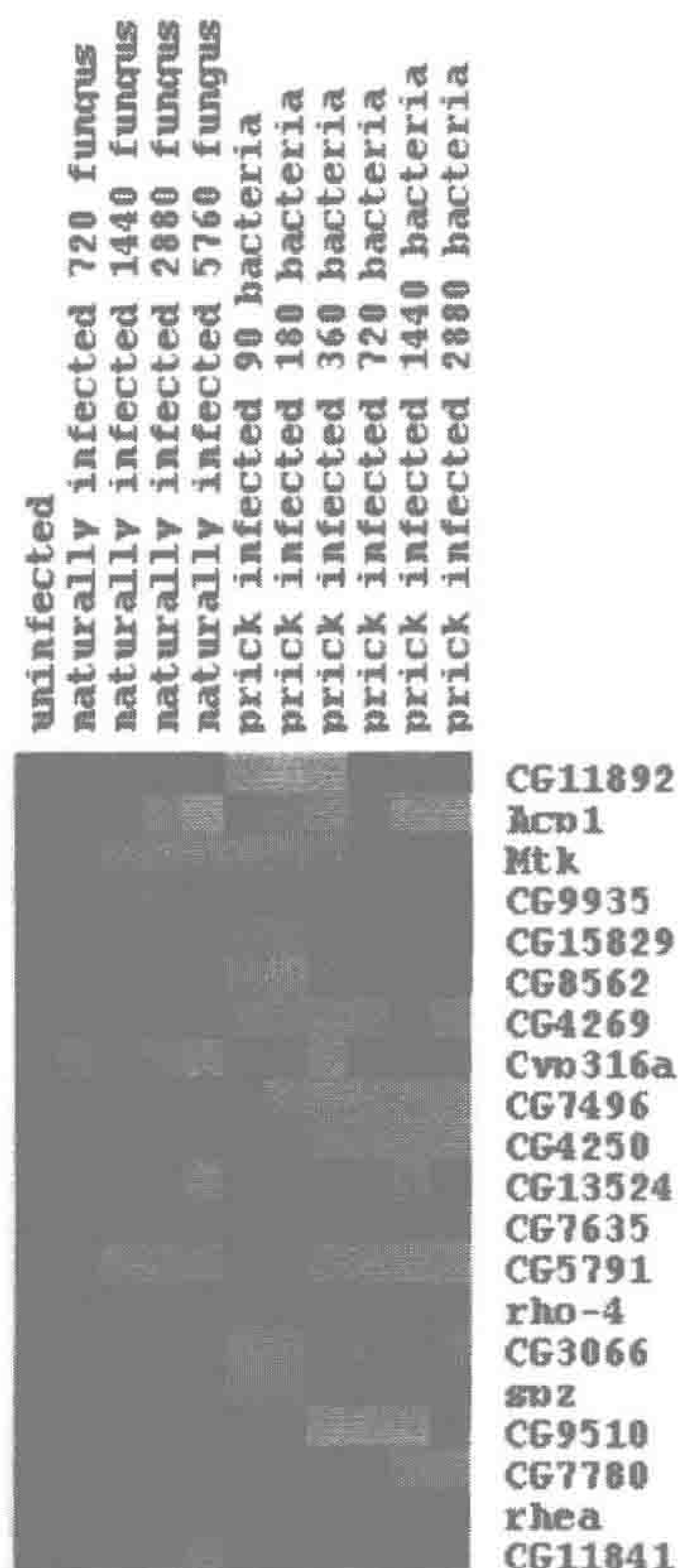
聚类方法的另一个重要的用途是给生物学样本分类，其目的在于识别不同的子集。关于这种类型分析好的例子有肿瘤样本的分类（Alizadeh et al. 2000；Perou et al. 2000）和伴随特定的基因组突变而引起的转录变化（Hughes et al. 2000）。有监督和无监督方法都可用于数据的聚类分析。

这里介绍了基因表达数据聚类的四个方面：①微阵列数据结构；②Cluster 和 Tree View 软件包；③聚类的类型和数学原理；④数据的调整和过滤。

微阵列数据结构概述

表达空间

正如一个物体在三维空间中的位置可以用 x-y-z 坐标来表示一样，根据表达值所提供的基因在表达空间的位置，一个基因的向量可以表示为坐标。例如，设想某个实验中使用了三个阵列。每个阵列样点将会有三个对数比值，分别来自每一个阵列。通过对表达值的作图，将基因在三个阵列上的表达值分别在 x 轴、y 轴和 z 轴上作图，任何人可以容易地将一个基因在三维空间中表示出来。这就得到了三维空间中的一个向量，表示了基因整体上的表达谱。同样方法，有 n 个阵列的实验，其基因表达的向量在 n 维空间作图。该空间被称作表达空间。表达谱相似的基因在表达空间中具有相似的向量。样本



571

图 7-39　一幅基因表达图

图中表示了在 11 种实验条件下对 22 种基因表达的测量。每个列的上方是对实验的描述，而基因的标识符（*Drosophila* CG 或者基因名称）在每一行的右侧。(红）与未被感染样本相比表达量升高；(绿）表达下降。颜色的强度对应于改变的幅度，明亮的颜色表明与未被感染样本之间有较大的差别（见彩图）

也可以用向量来表示，此时每一维对应于阵列上的每一个样点。来源相似的样本（相似来源的组织或细胞系）预期会有相似的基因表达谱，并且它们的向量将会与 n 维空间中相对比较接近的点有关。

聚类是一个广泛性的术语，可以应用在基于潜在的相似性来组织数据的许多种方法中。有许多种不同类型的聚类方法。常用的方法有层次式聚类、k-均值聚类和自组织图。这些方法具有许多共同的特性，特别是度量样本或基因在表达空间中的相关性的距离测度。距离测度是计算基因相互之间在表达空间中“距离”的一种方法。不同聚类算法的输出在很大程度上依赖于所使用的距离测度的类型，而不是其他的因素。不同距离测度的属性将在下面详细地研究。

基因表达图

由于一个典型的微阵列实验具有数量庞大的数据，为了能够容易地表现出数据内在的含义，必须对输出结果进行简明的显示。文献中基因表达图有许多名称，但是它们的基本原理都是 Eisen 等（1998）所介绍的。这些图一般将数据表示为一幅图像，其中每个基因为一行，每个实验为一列（图 7-39)。通过为逐渐偏离 0 的对数比值指派逐渐变亮的渐变颜色来显示基因表达比值。一个有用的惯例是用红色来表示升高的基因表达，而降低的基因表达用绿色表示［经过修改的热图 (heat map)］，虽然另一种蓝（降低）和黄（升高）的方案也很常用。

聚类和可视化工具

应用于微阵列数据分析两个最初的软件（Cluster 和 Tree View）由 Eisen 博士编写。这两个工具都可以从 Eisen 的网站（http://rana.lbl.gov）下载，进行直接安装后可以在 Windows NT/2000/XP 操作系统下运行。也可以得到具有不同功能的一些其他免费软件包，这里不再评论。

这里使用了选自 De Gregorio 等（2001）提供的基因表达数据的一个小的子集（400 个基因）来说明对微阵列数据进行聚类所涉及的方法。这些数据通过用真菌或细菌病原体感染果蝇（*Drosophila*)，然后用微阵列分析基因表达情况而得到，可以从 http://www.fruitfly.org/expression/immunity 在线检索，网站上表格中的所有比较都是

针对于未感染的样本（0 时刻）。

Cluster

Cluster 是一个用于对微阵列数据聚类的 Windows 应用程序。它可以上载、过滤和处理微阵列数据，以便进行层次聚类和 k-均值聚类，用于生成自组织图（SOM）和进行主成分分析（PCA）。

全部这些功能都是通过主窗口中的菜单来控制的（图 7-40 中所示）。Cluster 具有极好的手册（http://rana.lbl.gov/manuals/ClusterTreeView.pdf），其中介绍了它的各种选项应该如何应用。很多相关的性能也会在下面的聚类类型和数据问题中加以介绍。

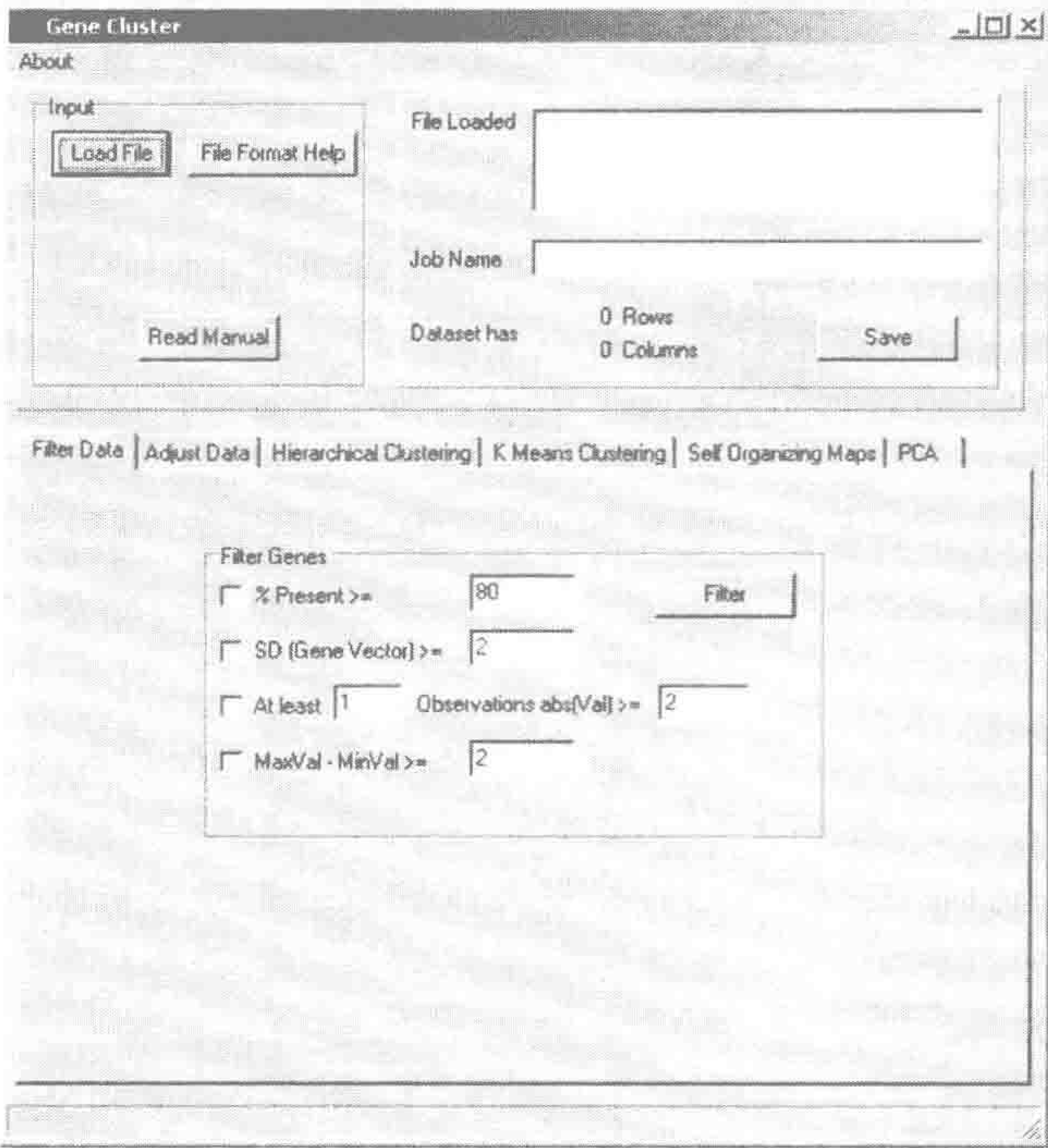


图 7-40　Cluster 应用程序

572

Cluster（http://rana.lbl.gov/）是一个组织基因表达数据的简单应用程序。用户可以通过“Load File”按钮来选择数据文件，“Save”按钮将当前的数据集合保存到一个文件中，包含了在 Cluster 中所作的修改。六个选项卡对 Cluster 的分析功能进行了组织：Filter Data（过滤数据，图中显示）、Adjust Data（调整数据）、Hierarchical Clustering（层次聚类）、k Means Clustering（k-均值聚类）、Self Organizing Maps（自组织图）和 PCA（主成分分析）。文中讨论了前三项，剩下的在 Cluster 手册中有说明

Tree View

直到现在，关于 Tree View 的说明文档还没有像 Cluster 的文档那样完备，因此在这里对它的性能进行比较详细的介绍。从根本上看，Tree View 是微阵列数据的一种可

视化软件包，它可以绘制基因表达图，生成适合输出的图像以及提供与外部基因组数据库的连接。

Tree View 包括两个主要的组成部分：简略视图和缩放视图，都用于显示基因表达图（图 7-41）。另外，缩放视图还显示对基因与实验的描述，而简略视图则显示了系统树图（dendrogram）。系统树图是一种将各个节点连接在一起的树状结构，每个分支点的长度表示对应的两个节点间的相似程度。系统树图被广泛用来表示系统发生关系。

Tree View 中的缩略图是研究基因表达聚类结构的一个非常有效的工具。因为系统树图是可以点击的，用户可以选择单独的分支，然后选中的部分会显示在缩放视图中。

Tree View 能够打开两种类型的文件，也就是纯文本文件和聚类数据文件，这两种文件分别通过后缀. txt 和. cdt 来确定。Cluster 使用手册中对文件结构有详细的描述。简而言之，Tree View 认为文件的起始两列是基因叙词（一个唯一的标识符和一个描述性字符串），并且起始两行是实验叙词。另外还有一些行和列是由 Cluster 加入的，其目的在相关的文档中有说明。假使阵列树（. atr）文件和基因树（. gtr）文件与. cdt 文件处在相同的目录下，Tree View 将为实验和基因产生树结构。如果使用一个一般的文本文件，就不会画出树图，但是全部的其他功能仍然可用。

Tree View 中有一些通用的控制机制。点击和拖拽面板之间的边界可以调整树图、基因描述、实验描述或缩略视图的宽度。但是，如果部分显示过大或者过小，在当前运行的程序（Version 1. 5，April 2000）中，调整大小可能不能正常地工作。Setting 菜单中 Options 表格允许用户调整图像的尺寸、颜色强度和色标。最后，通过使用“Find”命令在基因描述信息中查找某个特定的词，可以很容易地确定处在一个类别之中的单个基因。

574

Tree View 最有用的特性之一是它能够很容易地访问外部支持 Web 的数据库（web-enabled database）。当程序启动之后，它自动地在安装目录下搜索名为 URL-strings. txt 的文件。这个文件中包含一个针对多种数据类型的规范的 URL 集合。在缺省情况下，它提供了从酵母 ORF 标识符到斯坦福基因数据库（stanford genome data-base）、从小鼠 RIKEN 克隆到 RIKEN 数据库的关联，以及映象和登陆 Standford SOURCE 页面的通用登录号（general accession number）。为实现这一功能，. cdt 文件的第一列必须标明一个与 URLstrings. txt 文件中的起始条目相同的字符串，文件中该行剩余部分包含一个“＝”符号和紧随其后的 URL。唯一的 ID 必须插入在数据库所认为的查询位置处。下面给出的两个例子是 Tree View 中默认的：

YORF＝http://genome-www. Stanford. edu/cgi-bin/dbrun/SacchDB?find＋Locus＋%22UNIQID%22

MCLID＝http://genome. rtc. riken. go. jp/cgi-bin/getseq?g＋R＋UNIQID

YORF 指明了一个关联到 *Saccharomyces* genome database（SGD）的酵母可读框（yeast open reading frame）。第二个例子是针对 RIKEN 克隆的。当基因被点击后，这些例子当中的 UNIQID 项将被该基因的. cdt 文件中标明为 YORF 列中的字符串所取代。例如，假使 YORF 为 YLR452C 的酵母基因 *SST2* 被选中，检索将被替换成：

http://genome-www. standford. edu/cgi-bin/dbrun/SacchDB? find ＋ Locus ＋%22YLR452C%22

图 7-41 Tree View 应用程序

573

A. Tree View 主窗口。黄色方框表示简略视图中被选中的区域，对应的树中部分（如果存在的话）显示为高亮度紫色。屏幕左侧显示当前所加载数据的简略视图，而右边显示了当前选中数据的缩放视图。两幅视图之间的分隔边界可以改变尺寸，分配用于树图显示，基因描述和实验描述的区域也是如此。B. Tree View 的 File 菜单。“Load” 允许用户选择一个数据集合（. cdt 或. txt）。其余的选项用来将树图，简略视图，或者缩放视图保存为 PostScript（. ps）格式或者 Windows 位图（. bmp）。C. Setting 菜单。点击 Options 可以出现一个面板，允许用户更改显示特性：简略视图和缩放视图的行和列的大小，以及对比度（使用最亮颜色时的对应值）；例如，当该值为 2 且调色板（palate）是红/绿时，所有对数比值大于 2 的测量都表示为最亮的红色，而对数比值小于 −2 的用最亮的绿色表示）。Color 面板允许调色板从默认红/绿的方案调整为用户自选的（见彩图）

这时的输出将提供在 SGD 中关于 *SST2* 的结果。不幸的是，该工具受限于不同数据库的语法不一致。然而，通过在感兴趣的数据库中使用不同的基因进行 Web 查询测

试，应该可以揭示出用来转变为 URLstrings. txt 文件的适宜的数据库查询结构。使用一个能为数据库所理解的标识符是必需的（例如，对于 SGD，使用 ORF 名称如 YAL034W）。

Tree View 可以输出能供给期刊出版用的高质量图像。使用 File 菜单中相应的保存命令，用户可将他们的聚类结果存储为 postscript 或 bitmap 图像。将图像保存为 postscipt 格式并在应用程序如 Adobe Illustrator 中进行修改，可以在最大限度上控制图像的质量。

575

聚类类型

Cluster 中可以进行四种分析：层次聚类、*k*-均值聚类、自组织图（SOM）和主成分分析法（PCA）。Part Ⅵ着重介绍层次聚类。在层次聚类中，建立了每个基因（或实验）之间的关系，不管这些关系是否存在着生物学上的意义。在这种方法中构造了一棵聚类树，其中在一系列样本中表现出相似表达谱的那些基因在位置上靠得很近，而表达谱不同的基因彼此分开较远。聚类算法在生物学上常被用来构造系统发生树。大多数的聚类方法是凝聚的（agglomerative），也就是说，聚类树是从底部开始构造的。首先将最相似的基因归并在一起，然后逐步向上，直到所有的分支都归并为一个类。有些层次聚类工具提供分裂式（divisive）聚类，其中数据被不断地分割为更小的集合（Alon et al. 1999）。层次聚类工具通常是确定性的，也就是说，对于相同的输入数据它们总是生成同样的聚类树，虽然应该注意到输入数据集的细微改变可能会导致输出结果中发生重大的差别。凝聚式层次聚类法是一个逐步的过程，从计算基因和（或）实验间的全部配对距离关系开始。最相近的成对值被合并形成一个节点（node），这个步骤重复进行，比较基因（或新的类别）直到所有的类别归并在一起。

随着数据集复杂性的增加，许多层次聚类方法受到计算限制的约束。计算时间与基因数量的平方成比例，而与实验的数量呈线性关系。例如，假定使用 1 GHz 的台式计算机，处理一个小的数据集（约 1000 个基因×约 50 个实验）只要花费几分钟，但一个大的数据集（约 14 000 个基因×约 300 个实验）则需要 2 h，而一个更为庞大的数据集（约 40 000 个基因×1000 个实验）将要花费约 50 h。依赖于程序的实现，随机存取内存的不足会是可能遭遇的最后一个绊脚石，这将限制用户对大型数据集的聚类。

距离测度

距离测度是用来确定一对基因（或实验）表达谱相似性的一种计算公式。虽然 Cluster 和其他工具提供了许多种距离测度，通常最好还是使用 Pearson 相关系数，因为该测度对真实比值不敏感，而是基因间变化的相对量。Pearson 相关系数的计算是用两组数据的协方差除以每组数据的标准差之积（等式 1）。当参考样本与测试样本在生物上相关时，如在时间序列实验当中，使用零时刻样本作为所有其他时刻样本的参考，平均基因表达假定为 0 的非中心化 Pearson 相关系数（等式 2）非常有用，

$$r = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{X_i - \bar{x}}{\sigma_x}\right]\left[\frac{Y_i - \bar{Y}}{\sigma_y}\right] \tag{1}$$

$$r = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{X_i}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_i)^2}}\right]\left[\frac{Y_i}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(Y_i)^2}}\right] \tag{2}$$

576

Cluster 中还包括了在文献中使用的许多其他距离测度。欧几里得（Euclidian distance）距离，是毕达哥拉斯（Pythagorean）定理的一种推广，等于每对基因各坐标差值平方和的平方根（等式 3）。

$$\text{距离} = \sqrt{\left(\sum_{i=1}^{N}\right)(X_i - Y_i)^2} \tag{3}$$

斯皮尔曼秩相关（Spearman rank correlation，*rho*）和 Kendall's *tau* 忽略了实际的基因表达值，替之以它们的秩来确定相似性。人们还提出了一些其他的测度并用于微阵列数据的分析。来自 Heyer 等（1999）的一个例子是“刀切法相关系数（jackknife correlation）”，其中去除了可能代表了错误的极端值。直到此时，Cluster 当中还没有包括刀切法相关系数，但在以后的版本中可能会加进去。

另一个受用户控制的变量是连接（linkage），它是描述节点如何结合的一组规则。建议使用 Cluster 中的平均连接聚类法（average linkage clustering）选项，因为该方法对噪声的敏感性最低。

图 7-42 是在 6 个“阵列”上测量 4 个“基因”得到的数据的折线图（line graph），而表 7-8 中是根据上面介绍的各种相关性度量经过计算得到的每对结果的相似性评分。表中的前三个距离测度量（中心化 Pearson、非中心化 Pearson 和 Spearman 秩）的值都位于 −1～1（值越大表明相似性越强），而 Euclidian 距离的值可以从零到无穷大（值越小表示相似性越强）。在这组数据中可以看到几个非常大的差异，特别是当用中心化和非中心化 Pearson 相关系数来计算时，“基因” A 和 B 之间从完全相关变化到有限的负

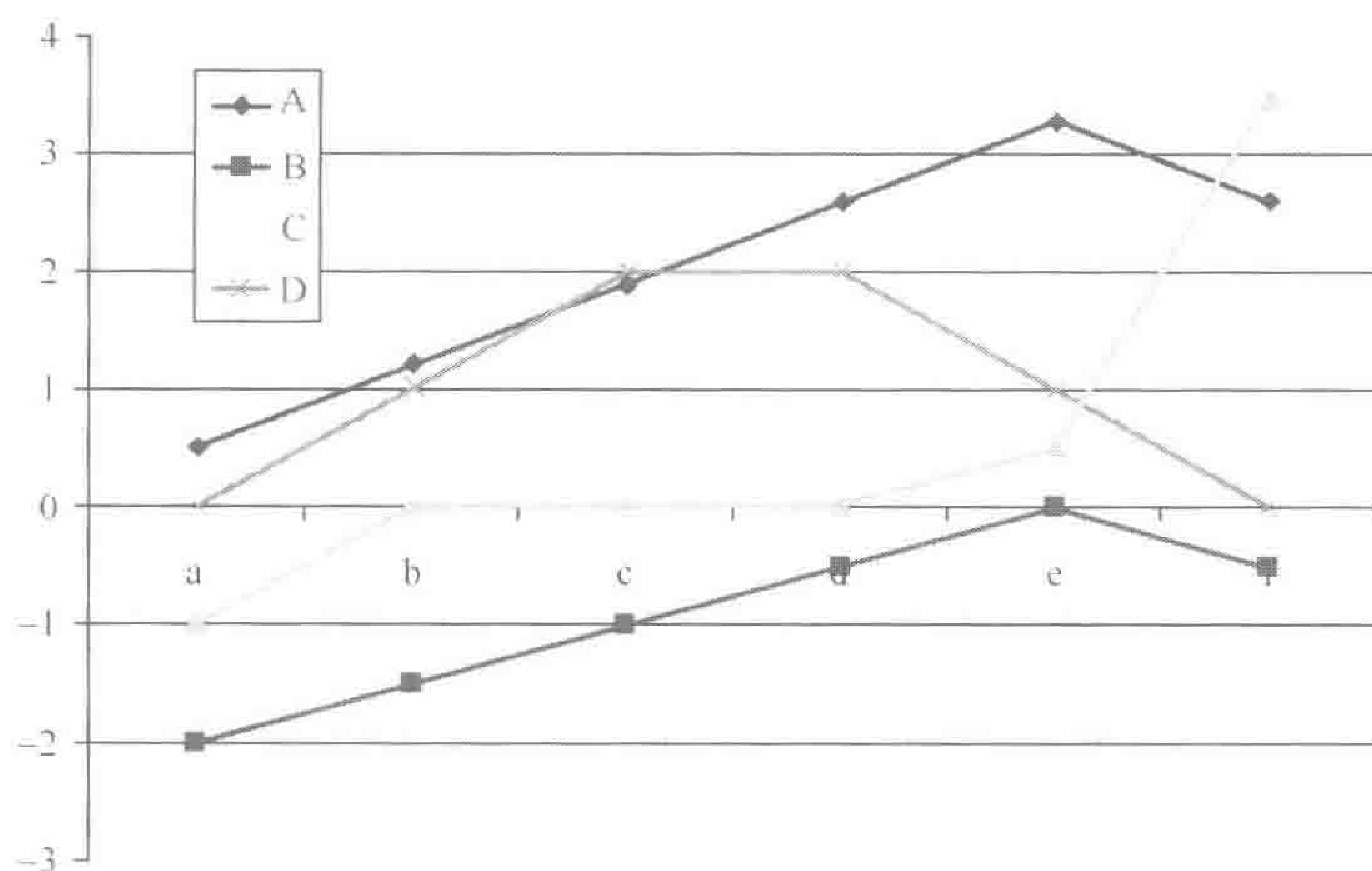


图 7-42　四个基因样本的对数比值作图

四个假想基因（A～D）在 6 种实验条件（a～f）下表达的对数比值的折线图。注意到基因 A 和基因 B 的曲线形状很相似，并且在计算中心化 Pearson 相关系数值时两者是相关的，而使用非中心化的测度时却是不相关的（请见表 7-8 和正文）

相关，这是在计算相关时关于中心化影响一个极好的例子。如果这些数据是相对于时刻零的时间序列，A 和 B 将不被认为是相似的，但是如果它们是与生物学上无关样本进行比较的肿瘤样本，那么这些基因的行为将可能会被认为是相似的。

表 7-8　不同测度下各对基因的距离

	中心化 Pearson[a]	非中心化 Pearson[a]	Spearman 秩[a]	欧氏距离[b]
AB	1.00	−0.48	1.00	7.22
AC	0.53	0.51	0.75	4.77
AD	0.30	0.78	0.30	3.57
BC	0.53	0.02	0.75	4.56
BD	0.30	−0.44	0.30	5.17
CD	−0.36	0.04	−0.22	4.74

a 较大的数字表明关系较密切的成对基因（取值范围−1～1）；b 较小的数字表明关系较密切的成对基因（取值范围 0～∞）。

数据问题

比值

通过标记的 cDNA 与单个微阵列的竞争性杂交，玻片微阵列通常用于测量两个样本中 RNA 的相对丰度。这类实验的标准输出是微阵列上每个样点在每个通道中信号的比值。输出是不对称的这一事实，也就是说，两倍变化的比值可以是 2（上调）或 0.5（下调），导致了分析上的困难，因为根据方向的不同，大多数距离测度对比值变化的处理是不同的。因此，将比值数据转化成一种对方向变化不敏感的形式是很关键的。

有两种常用的方法可以解决这个问题：①将全部比值转化为对数（以 2 为底的对数比较方便，但所有的底数都是等效的）；②不是计算两个值之比，而是计算相对于基线情形的倍数变化，并表示为(I1-I2)/(I1或 I2中较小的数值)，其中 I1 是第一次测量的强度，I2 是第二次测量的强度。

将数据表示为对数比值和倍数变化。这两者相比主要的区别是使用对数比值时高的数值被压缩了，而在倍数变化中则不是这样。对数比值压缩带来的代价是通过对较小比值的相对放大增加了较小比值的重要性。对数比值的一个优点是，在微阵列实验中对数比值的分布通常是近似于正态分布，因为大多数类型数据所具备的一个共性是对数变换使它们正态化，即使它们起初不是这样的。虽然这个特性在其他类型的分析和数据处理中发挥着作用，但是对于聚类却不是关键的。实际应用中采用对数比值还是倍数变化，更多地依赖于数据集的特性而不是其他的因素。图 7-43 中显示了分别使用对数比值和倍数变化进行聚类而得到的结果。

数据的中心化和标准化

在大多数情况下，扫描得到的数据在进一步的分析之前需要作变换和标准化处理。

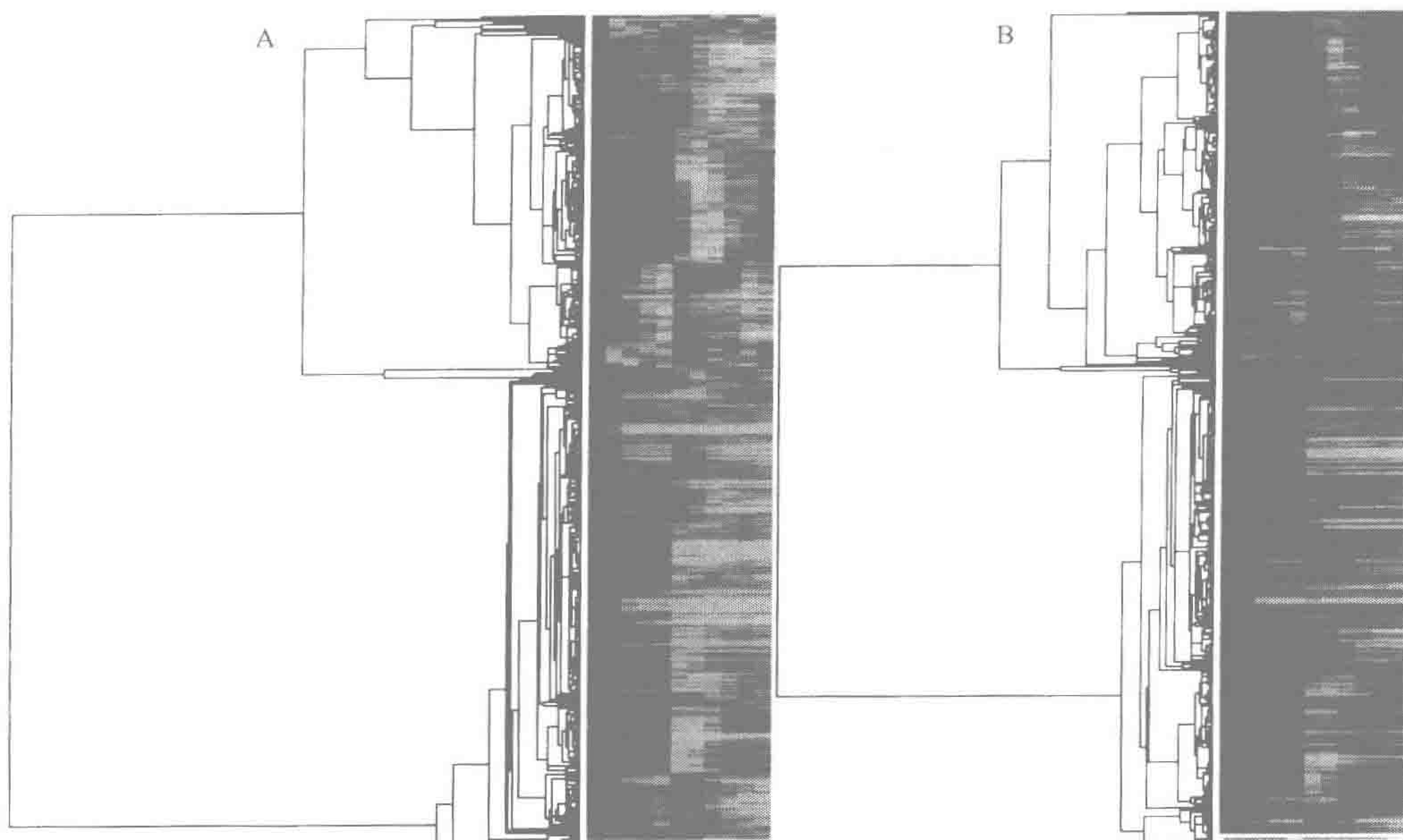


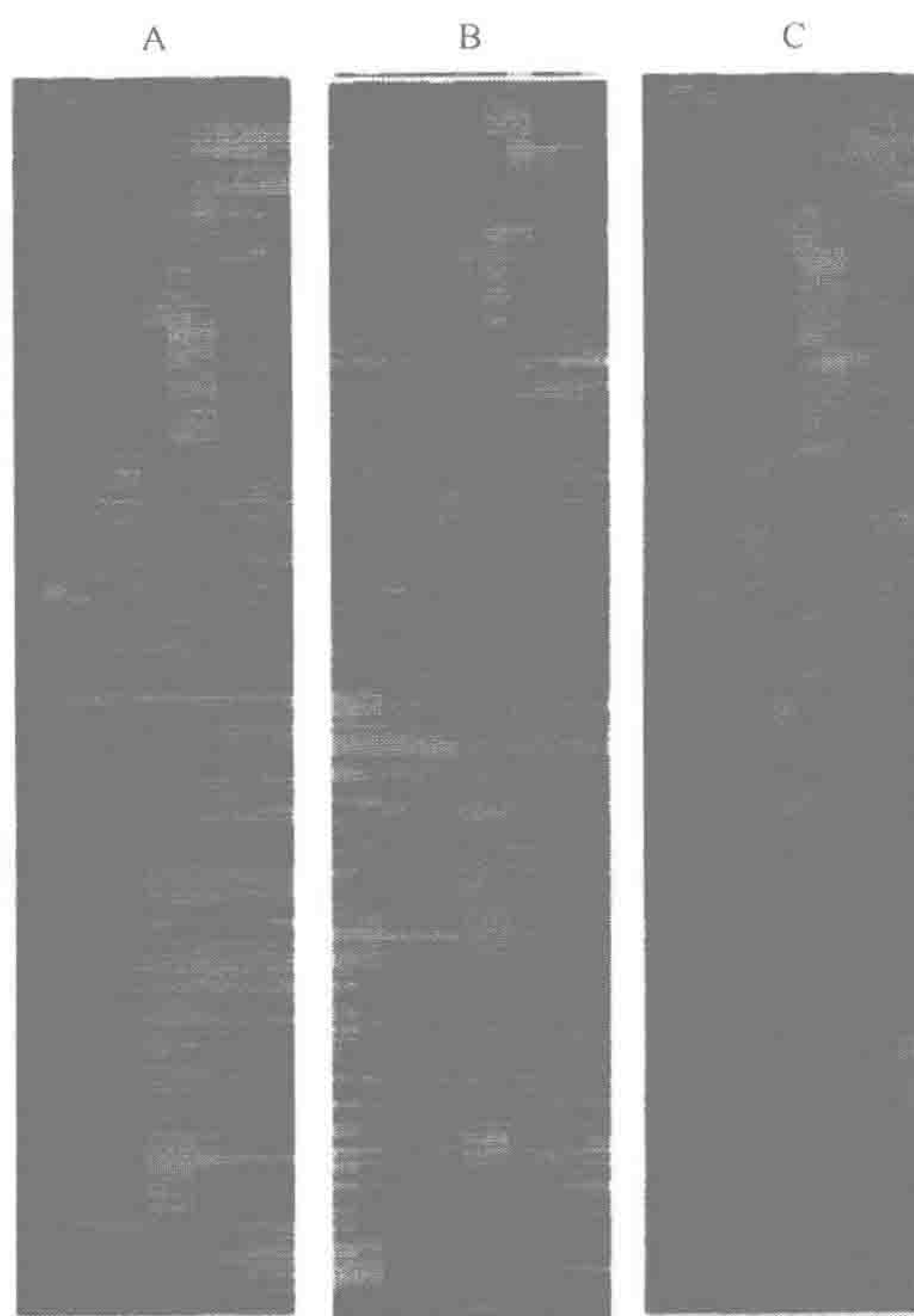
图 7-43　对数比值和倍数变化的比较

对数比值形式（A）和倍数变化形式（B）的免疫应答数据的层次聚类结果。每个小格（图中的每一个矩形）的色彩强度与其值成正比。要注意根据倍数变化作图时小格为最大强度。但是，数据聚类结果的差别却很少。数据被分裂为两个主要的分支（抑制的或诱导的基因），而且在根分支下面的结构相对稀少（见彩图）

578

即使数据已经被转换成对数比值（或者一些等效的因子），仍然可能会需要额外的数据处理。参考样本的表达谱预期和测试样本之间没有极大的差别时，可以假定平均对数比值接近于 0，并且其中一个样本的数据可以用标准化因子来进行调整，从而使得对数比值为 0。如果数据集是使用与测试样本存在很大表达谱差异的生物学上无关参考生成的，对于实验集中的每个基因，平均对数比值（或者等效的）未必为 0，而且事实上有可能远离于 0。

这时如果采用非中心化的 Pearson 相关测度，在一系列样本之中表达谱相同但是在参考样本中不同的两个基因，在应用简单的对数比值强度时将不能聚类在一起。从一个基因的每个对数比值数据点之中减去该基因的平均对数比值可以解决这一问题（“均值中心化”）。在这里作为例子的数据集当中，数据显示为“0 中心化”最为合理，这时从每个时间点的对数比值中减去 0 时间刻的对数比值。图 7-44 的 A 和 B 分别给出了每个基因当数据为 0 中心化和平均值中心化时聚类输出结果的比较。中心化的 Pearson 相关也可以应用于这种类型的分析，但是这样做将会导致异常的显示图案，其中主要为红色的基因（0 对数比值之上）在聚类树结构中将和显示为绿色的基因相邻。正如从表 7-8 中可以看出来的，基因 A 和 B 的中心化 Pearson 相关系数为 1，而非中心化 Pearson 相关系数为 −0.48。如果对这些基因采用“均值中心化”后的数据作图，它们将是一致的，但是在“非中心化情况下”，它们显然差异很大。另一个也可以使用的非常类似的变换采用对数比值的中值进行（“中值中心化”）。中心化同样也可以应用在“实验”轴上。



579

图 7-44　经过中心化和标准化的数据

时刻 0 中心化的（A），均值中心化的（B）和经过标准化的时刻 0 中心化的（C）免疫应答的对数比值数据。每组一系列数据均进行层次聚类（树结构移去）。均值中心化的数据（B）与 0 中心化的数据有着截然不同的外观，虽然聚类结构还算类似（但离一致还很远）。标准化的数据显示了相似的外观但强度更弱，这使得更多的变化可以被观察到，而 A 中的许多基因在所有实验中显示出最高的强度值从而掩盖了它们的变化（见彩图）

在一些场合下，对基因表达进行标准化是恰当的。标准化和中心化是有区别的，因为它是将基因的表达向量的相对大小调整为 1，但是并不改变数据的形状。标准化常用在自组织影射图或 k-均值聚类当中以防止表达变化大的基因在输出结果中占据主导。当数据没有被标准化时，一些基因的极端值不容易显示出来，这时进行数据标准化也会使得对这些基因在基因表达中变化的查看变得容易。在这种情况下的标准化是将基因的表达向量的大小设置为 1（单位向量），方法是用向量除以它的大小（向量的长度）。测试数据集经过标准化和聚类后的输出结果见图 7-44C 中。

数据冗余性

由于一些有意识的或者事先未预料的原因，产生的数据集常常是部分冗余。假定结果是能够独立验证的，那么数据冗余常常是一件好事。

580

然而，由于会夸大一些测量的值，部分冗余会改变（有时是完全改变）聚类方法的结果。阵列中一些样点量的冗余可能会导致实验聚类时产生所不希望的偏差，而样本的冗余则会引起基因聚类时的偏差。例如，假设一种真菌感染时间序列被测量了两次，但是一个败血症损伤时间序列只检测了一次，这时数据将是部分冗余的，并且潜在性地偏向于测量次数多的真菌序列。

两种方法可以解决与数据冗余有关的问题，虽然它们都要求花费些努力以保证具有客观性。Cluster 具有自动加权的特性以降低非常相似数据系列的重要性。但是，这个特性并不是像它的名字看上去那样地自动。实际上，用户必须输入一个他们期望用来反映数据采集偏差的相关值和一个用于加权的幂函数。更完整的解释请见 Cluster 使用手册。另一种可以选择的方法是，用户可以在数据文件中标注出基因和实验的相对权重。通常，完全冗余数据的信息含量与观察次数的平方根成正比，因此 9 次冗余实验的信息含量等于 3 次非冗余实验。9 次重复实验中的每一次都要加权为一次非冗余实验的 1/3。部分冗余数据由于不能进行平均而更加难以处理，因此在这种情况下最好采用自动加权特性。

噪声

微阵列数据测量容易产生噪声（定义为对同一对象的重复测量中观察到的差异）。这个问题有两个主要的解决办法，但是当前在 Cluster 中都还没有被包括进去。由于经验证据表明低强度的样点更加可能是噪声性的，去除这些样点是相当安全的选择。另一个降低噪声影响的选择是对重复测量进行平均处理。该方法的一种变化是数据的平滑处理，即在实验中的时间序列或者剂量响应中对相邻的点进行平均。实验平均和数据平滑都可以在电子表格软件，如 Excel 中很容易地实现。

降低数据复杂程度（过滤基因）

在两种主要的情况下需要降低数据的复杂性：去除噪声性数据和删除“不感兴趣的数据”。如上所述，Cluster 依赖于已经过滤的除去了每个测量中噪声的数据。对于单个基因，如果很大一部分的观察结果被过滤掉，那么很可能关于这个基因的所有测量都是不可信的，或者是观察到的表达谱不具备代表性。当然，假使基因的数据点留下来的少于一半，那么从数据集中去除这些数据可能是个好的主意。

减少要聚类基因的数量经常是有利的。有些减小数据集大小的理由很牵强（例如，大量基因会难以处理而且解释起来也困难），但是去除变化很小的基因除了只是方便外还有更好的证明。首先，保持不变的基因，如果数量很多，对聚类结果会有负面的影响。第二，展示很小的生物波动的信号常常被与测量有关的随机噪声遮蔽。

Cluster 中提供了三个独立的参数用于只将具有较大变化的基因包括在分析中：①最小标准偏差；②至少有 n 次测量的变化要大于某个最小值；③最小差异（最大值一最小值）。一般情况下最好只设置其中一个参数，而不是将它们结合起来使用，否则会导致不可预知的结果。

总结

层次聚类和基因表达图是当前微阵列分析中的主要分析工具。但是这决不意味着它们是最终的答案。特别是聚类，它应该是作为数据分析的起点来启发人们的思维，而表达图仅仅是一种简单的数据可视化方法。用户也应当寻找新的应用程序，可以同时进行探索性的发现和结构化的检索。虽然因为一些原因 Cluster 和 Tree View 成为

主要的分析工具；但是在使用时的参数设置正确的情况下，它们确实是一种有力的分析工具。

582

PART Ⅶ 多维尺度法和自组织图

Sujatha Panavally，Yidong Chen（国家人类基因组研究所，国立卫生院，贝塞斯达，马里兰州）

多维尺度法

对来自于一组实验的基因表达谱的研究需要有一个针对所有成对实验的相似性度量矩阵。尽管常用的聚类算法，如在 Part Ⅵ 中介绍的那些方法可以用来研究基因和样本之间的相互关系，还需要其他的模型用来对这些数据进行直观的显示（Khan et al. 1998；Bittner et al. 2000；Hendenfalk et al. 2001）。特别是在样本集有限的情况下，相关系数具有较宽的置信区间（Dougherty 2001）。多维尺度法（multidimensional scaling，MDS）分析技术是一种能够将相似矩阵的结构转换为简单几何图形的方法。该程序可以从 http://arrayanalysis.nih.gov 上公开获取。在 MDS 分析中，两个样本之间的不同点（通过基因表达谱来评价）越大，在图形中表示实验的点就分开得越远（Green and Rao 1972；Schiffman et al. 1981；Young 1987；Green et al. 1989；Borg and Groenen 1997；Cox and Cox 2000）。

一个基因表达数据集通常以矩阵的形式来表示，$\mathbf{X}=\{\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_M\}$，其中向量 $\mathbf{x}_i$ 表示第 i 次基因表达实验，在数据集中一共有 M 次实验。每次实验中包含了 n 个基因的表达水平或比值，表示为 $\mathbf{x}_i=\{\mathbf{x}_{i1},\mathbf{x}_{i2},\cdots,\mathbf{x}_{in}\}^T$，这里 T 表示矩阵的转置。矩阵 $\mathbf{X}$ 也可以在行的方向上来观察，这时提供了在所有样本之中的基因表达谱。样本可以根据时间或者剂量进行排列，或者它们也可以是无序的集合。PART Ⅶ 通过研究实验样本之间的关系来介绍 MDS 的概念和算法。

最常用的相似性测度是 Pearson 相关系数。给定一个矩阵 $\mathbf{X}$，相似矩阵 $\mathbf{S}=\{s_{ij}\}$，定义为

$$s_{ij}=\frac{\sigma_{ij}}{\sigma_i\sigma_j}=\frac{\sum_{k=1}^{n}(x_{ik}-\mu_i)(x_{jk}-\mu_j)}{\sqrt{\sum_{k=1}^{n}(x_{ik}-\mu_i)^2\sum_{k=1}^{n}(x_{jk}-\mu_j)^2}} \tag{1}$$

式中，μ_i 和 σ_i 分别是样本 i 的基因表达比值（通常要进行对数变换）的平均值和标准差。

一些其他的相似性测度，如 Spearman 相关系数或者标准化的欧几里得距离，同样适用于这里。通过 $\mathbf{D}=1-\mathbf{S}$，相似矩阵可以转换成距离矩阵 $\mathbf{D}=\{d_{ij}\}$。表 7-9 是一个典型的相似矩阵，其中对角线上的元素均为 0，并且上三角阵与下三角阵对称，每个元素的取值为 0～2。距离矩阵的大小为 $M\times M$，因为考虑的样本共有 M 个。

表 7-9 相似矩阵 S 的一部分，根据提供的数据样本产生

	UACC 383	UACC 457	UACC 3093	UACC 2534	M92 −001	A-375	UACC 502	M91 −054	UACC 1256	UACC 091
UACC383	1									
UACC457	0.60	1								
UACC3093	0.63	0.68	1							
UACC2534	0.56	0.51	0.59	1						
M92-001	0.53	0.46	0.54	0.59	1					
A-375	0.54	0.54	0.55	0.53	0.64	1				
UACC502	0.43	0.43	0.48	0.46	0.58	0.50	1			
M91-054	0.51	0.48	0.53	0.46	0.58	0.50	0.57	1		
UACC1256	0.49	0.51	0.51	0.52	0.59	0.53	0.60	0.54	1	
UACC091	0.52	0.44	0.50	0.47	0.53	0.52	0.51	0.59	0.58	1

Bittner 等（2000）。注意距离矩阵 $\boldsymbol{D}$ 等于 $1-\boldsymbol{S}$。

因此，通过使用距离矩阵 $\boldsymbol{D}$，我们有效地将系统的维数从 N（每个样本中总的基因数目）降低到 M，M 明显要小于 N。多维尺度法的目标是进一步降低矩阵的维数，从而得到具有最小误差的图形化模型。挑战在于要在三维空间中精确地表示高维的表达空间。一般的，在图形化模型中使用笛卡儿（Cartesian）坐标系，并用欧几里得距离度量点之间的距离。

为了实现从距离矩阵到图形化模型之间的映射，必须在一维、二维或者三维空间中寻找到一个目标矩阵 $\hat{\boldsymbol{X}}=\{\hat{x}_{ij}\}$。伴随的距离矩阵 $\hat{\boldsymbol{D}}=\{\hat{d}_{ij}\}$从 p 维空间中的欧几里得距离导出。

$$\hat{d}_{ij}=\sqrt{\sum_{k=1}^{p}(\hat{x}_{ik}-\hat{x}_{jk})^2} \tag{2}$$

使得应力（stress）T 对拟合优度（goodness of fit）的一种度量最小化。这里，我们使用的应力 T 的定义为：

$$T=\sqrt{\frac{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(d_{ij}-\hat{d}_{ij})^2}{\sum_{k=1}^{n-1}\sum_{j=i+1}^{n}d_{ij}^2}} \tag{3}$$

如同 Everitt 和 Dunn（1992）所建议的那样，如果能够在三维空间或者更低维的空间中找到一个可以被接受的精确表示，其中 T 从 $T=20\%$（差）、5%（好）到 2.5%（优秀），那么通过 MDS 的可视化将是用于洞察数据结构一个精确而且有价值的方法。

实现和应用

MDS 的实现是根据 Everitt 和 Dunn（1992）提出的方法。在经典的 MDS 设置中，给定全部 M 次实验中的成对距离 d_{ij}，在 M 维欧氏空间中 M 个点的坐标，$\boldsymbol{X}=(x_{ij})$可以通过寻找 $\boldsymbol{B}=\boldsymbol{XX}'$的特征向量而得到，

584 且矩阵 B 的元素$\{b_{ij}\}$，可以通过下式进行估计

$$b_{ij} = -(d_{ij}^2 - d_{i\cdot}^2 - d_{\cdot j}^2 + d_{\cdot\cdot}^2)/2 \quad (4)$$

式中，$d_{i\cdot}$、$d_{\cdot i}$和$d_{\cdot\cdot}$分别是d_{ij}^2在i，j以及i和j上同时求和的结果。通过在X中选取起始的k个主成分构成目标矩阵$\hat{\boldsymbol{X}}=(\hat{x}_{ij})$，且$i=1,\cdots,M$和$j=1,\cdots,k$，然后使用一些优化方法［例如，最速下降方法（steepest-descend method)］根据等式（3）进行优化，获得一个 MDS 模型。关于 MDS 的更多信息，请查阅这部分结束的文献。

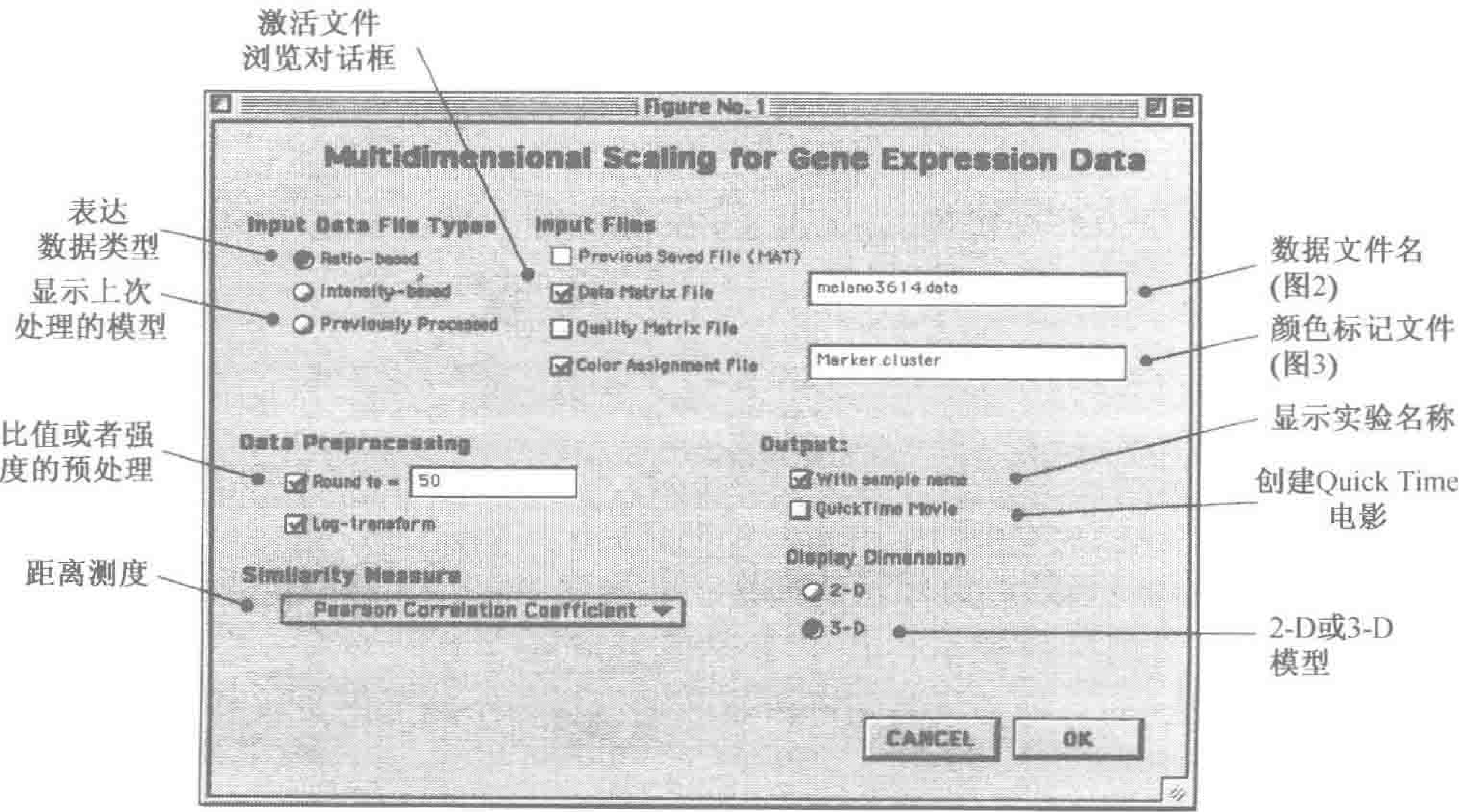
MATLAB 是一个数学分析和可视化的程序，作者的实验室中用它来实现一种 MDS 算法（Bittner et al. 2000）。MATLAB 的统计学工具箱对于这种 MDS 的实现是必需的。当前的 MDS 程序在用于 PowerPC Macintosh 计算机的 MATLAB v5.2 上测试通过，并且对其他类型计算机的支持在随同发布的 README 文件中有详细说明。以下软件包是 MDS 所必需的：

MATLAB 5.2(Macintosh)：http://www.mathworks.com/products/matlab/

Statistics Toolbox：http://www.mathworks.com/products/statistics/

统计学工具箱可以与 MATLAB 使用同一张光盘安装；然而，必须购买一个不同的注册码，这样在安装过程中才会出现该选项。在这两个软件包都被安装后，解压 MDS 文件到位于 MATLAB 应用文件夹下的新文件夹当中（命名为“MicroarrayTools”）。因此目录应该类似于...Matlab5.3 \ MicroarrayTools \ MDS \ 。

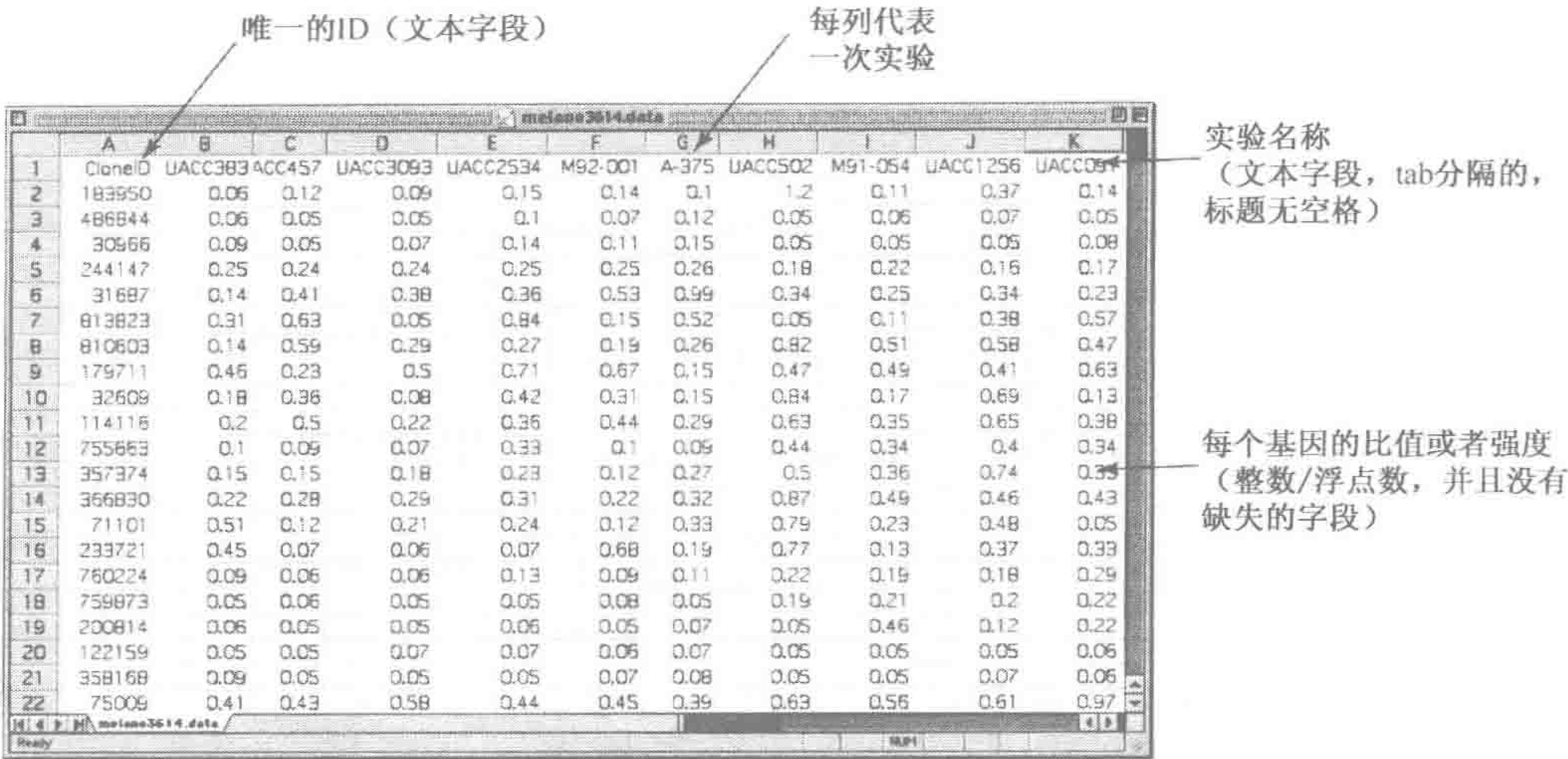
现在装载 MATLAB 应用程序，并从文件菜单中选择“add path”。找到 MDS 文件夹图标的位置，并点击“Append”按钮。要运行 MDS，在 MATLAB 主窗口的提示符处键入 mds _ dialog。应该出现如图 7-45 所示的对话框窗口。为了使用该程序，用户必须准备下列输入数据：



585 图 7-45　对话框窗口

• **数据矩阵（data matrix）**（可选择比值矩阵或者强度矩阵）。每次实验的表达比

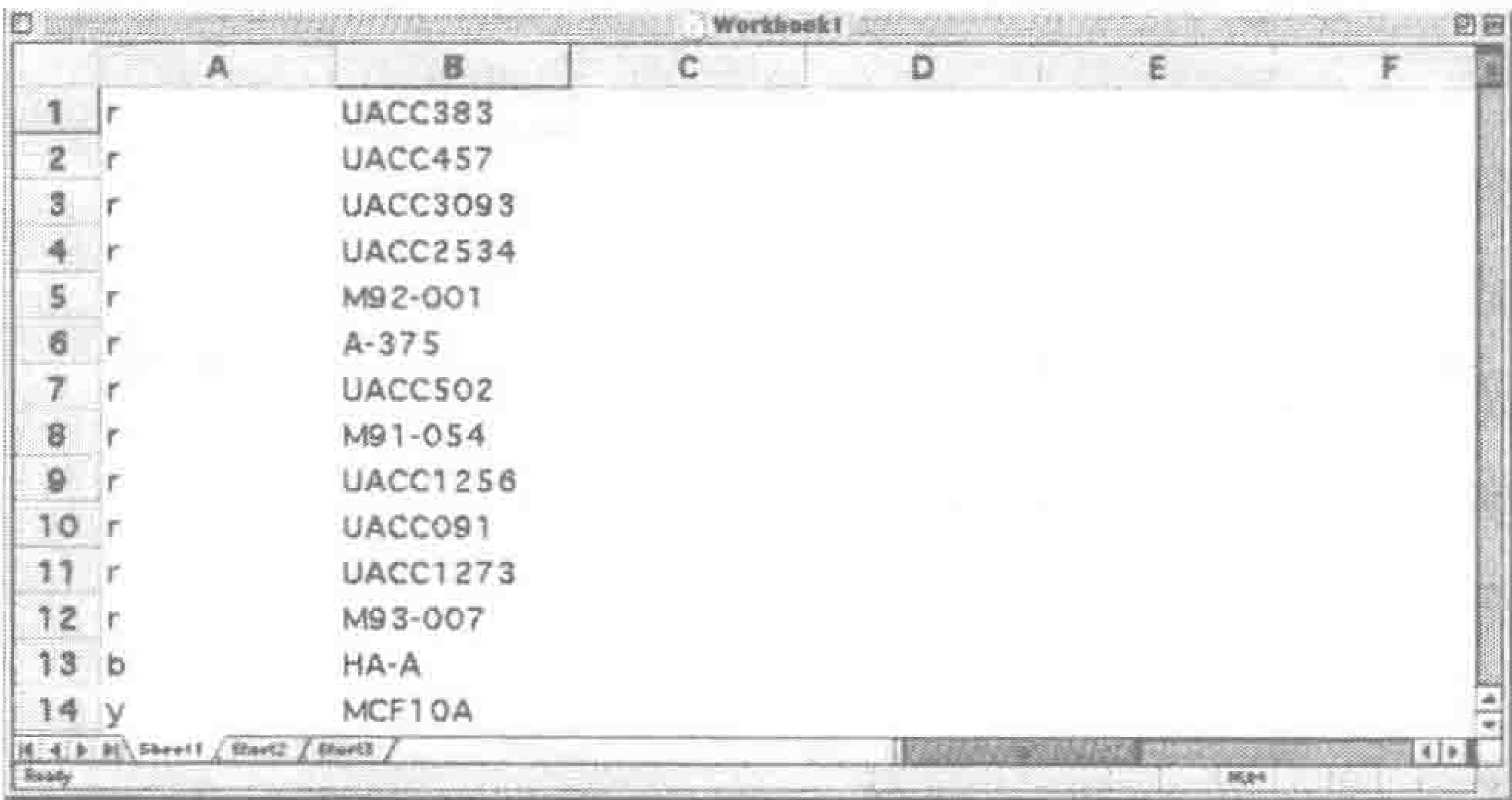
值（例如，样本 1 对样本 2 标准化的平均强度比值）或者强度按照列的方式排列，而每一行中给出了每个基因的表达数据。图 7-46 中给出了一个典型的数据表格。



	A	B	C	D	E	F	G	H	I	J	K
1	CloneID	UACC383	ACC457	UACC3093	UACC2534	M92-001	A-375	UACC502	M91-054	UACC1256	UACC09[illegible]
2	183950	0.06	0.12	0.09	0.15	0.14	0.1	1.2	0.11	0.37	0.14
3	486844	0.06	0.05	0.05	0.1	0.07	0.12	0.05	0.06	0.07	0.05
4	30966	0.09	0.05	0.07	0.14	0.11	0.15	0.05	0.05	0.05	0.08
5	244147	0.25	0.24	0.24	0.25	0.25	0.26	0.18	0.22	0.16	0.17
6	31687	0.14	0.41	0.38	0.36	0.53	0.99	0.34	0.25	0.34	0.23
7	813823	0.31	0.63	0.05	0.84	0.15	0.52	0.05	0.11	0.38	0.57
8	810603	0.14	0.59	0.29	0.27	0.19	0.26	0.82	0.51	0.58	0.47
9	179711	0.46	0.23	0.5	0.71	0.67	0.15	0.47	0.49	0.41	0.63
10	32609	0.18	0.36	0.08	0.42	0.31	0.15	0.84	0.17	0.69	0.13
11	114116	0.2	0.5	0.22	0.36	0.44	0.29	0.63	0.35	0.65	0.38
12	755663	0.1	0.09	0.07	0.33	0.1	0.09	0.44	0.34	0.4	0.34
13	357374	0.15	0.15	0.18	0.23	0.12	0.27	0.5	0.36	0.74	0.3[illegible]
14	366830	0.22	0.28	0.29	0.31	0.22	0.32	0.87	0.49	0.46	0.43
15	71101	0.51	0.12	0.21	0.24	0.12	0.33	0.79	0.23	0.48	0.05
16	233721	0.45	0.07	0.06	0.07	0.68	0.19	0.77	0.13	0.37	0.33
17	760224	0.09	0.06	0.06	0.13	0.09	0.11	0.22	0.19	0.18	0.29
18	759873	0.05	0.06	0.05	0.05	0.08	0.05	0.19	0.21	0.2	0.22
19	200814	0.06	0.05	0.05	0.06	0.05	0.07	0.05	0.46	0.12	0.22
20	122159	0.05	0.05	0.07	0.07	0.06	0.07	0.05	0.05	0.05	0.06
21	358168	0.09	0.05	0.05	0.05	0.07	0.08	0.05	0.05	0.07	0.06
22	75009	0.41	0.43	0.58	0.44	0.45	0.39	0.63	0.56	0.61	0.97

图 7-46　数据文件格式

- **质量矩阵**（**quality matrix**）（可选的）。该矩阵在结构上类似于数据矩阵。但是，质量矩阵中每个元素的取值范围为 0～1，表示从最差的测量质量（0）到最佳的测量质量（1）。MATLAB 的质量矩阵不能处理数据矩阵中的缺失记录，因此如果图像分析软件包提供了一个不完整的数据矩阵格式，那么在质量矩阵中必须把缺失记录转换为 0，而将剩余的元素设置为 1。
- **颜色分配文件**（**color assignment file**）（可选的）。这是一个简单的文本文件（图 7-47），其中为每个样本指派一个颜色编码来辅助输出数据的显示。在本书写作时，程序仅能选取颜色 r、g、b、m、y、k 和 w，分别代表红、绿、兰、青、紫红、黄、黑和白色。

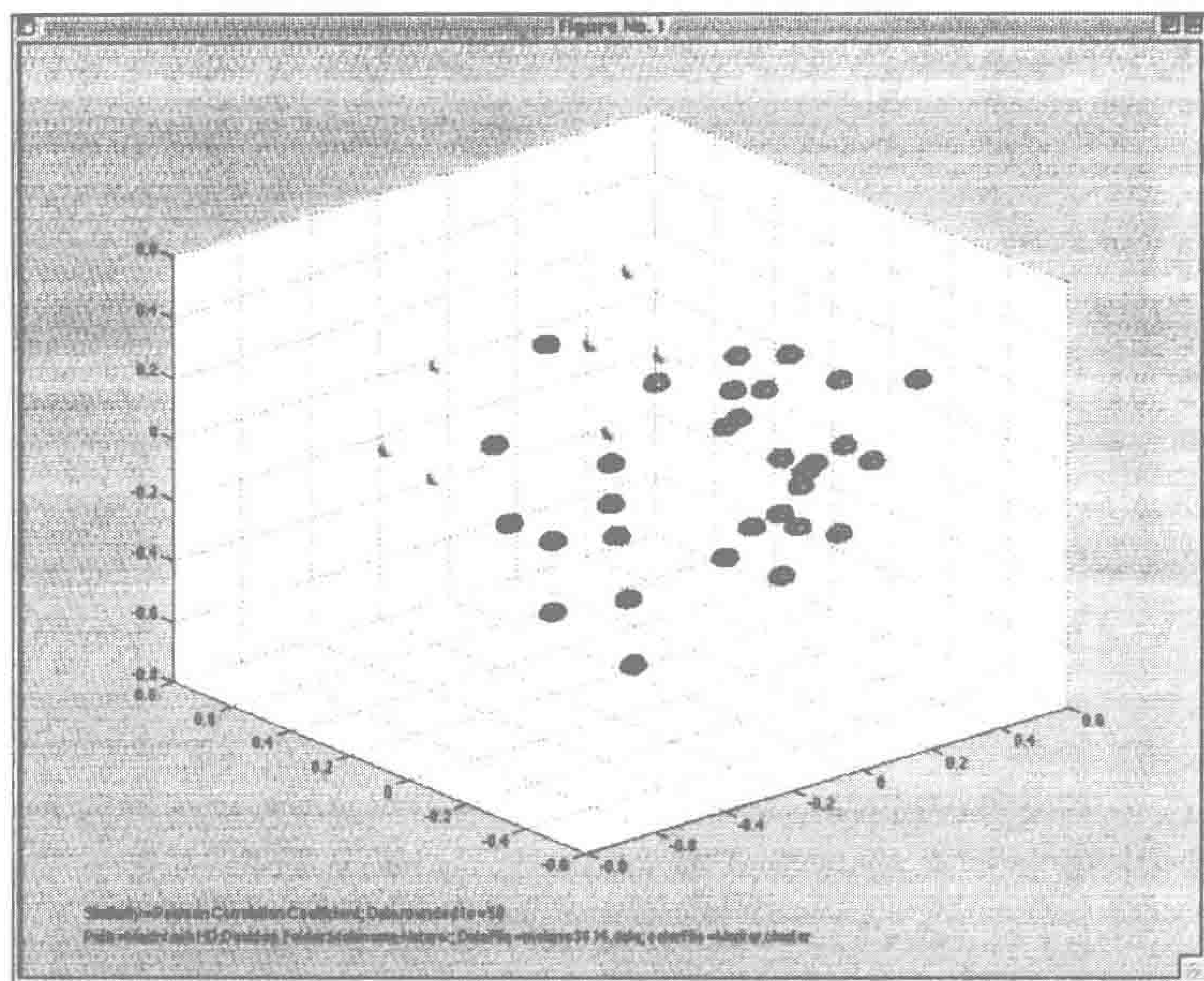


	A	B
1	r	UACC383
2	r	UACC457
3	r	UACC3093
4	r	UACC2534
5	r	M92-001
6	r	A-375
7	r	UACC502
8	r	M91-054
9	r	UACC1256
10	r	UACC091
11	r	UACC1273
12	r	M93-007
13	b	HA-A
14	y	MCF10A

图 7-47　颜色分配格式。第一列必须是颜色编码

用户可以选择一些预处理步骤，如比值的舍入（所有大于 50 或者小于 1/50 的比值将分别被舍入为 50 或 1/50）和数据的对数转换。一些距离测度以下拉式菜单的方式给出，而且用户能够选择最终的视图为带有或没有实验名称的二维或者三维图形，或者是用 QuickTime 电影作为最终显示。

图 7-48 中是一个典型的 MDS 显示。用户能够通过按住鼠标键，移动数据集来旋转坐标系统。



586

图 7-48　典型的 MDS 显示方式。三维图形显示

自组织图

自组织图（self-organizing map，SOM），有时也被称为拓扑有序映射或者 Kohonen 自组织特征映射（Kohonen 1995；Oja and Kaski 1999），是与多维尺度法密切相关的一种方法。该程序可以在 http://arrayanalysis. nih. gov 上公开获取。该方法的目的是用目标空间中的点来表示源空间（跨越基因表达谱的原始的多维空间）中的所有数据点，使得距离和邻近关系被尽可能多地保留下来。目标空间可以是一维的直线或者二维的网格。SOM 还可以解释为神经网络的一种特殊情形，其中使用了一种非监督的竞争性学习算法来对高维数据集进行可视化和解释。映射图中包含了一层节点或“神经元”，这里的每个节点有一个由特征构成的多维模型向量与之关联。映射图尝试使用有限的节点集合作为模型来表示所有的观察结果，并且节点的组织方式要使得在网格空间中相似的节点之间互相邻近，而不相似的节点则互相远离。

SOM 通过竞争性学习来训练。当输入向量提交给网络时，对于 SOM 中那些所包含向量与输入模式最为相似的节点，被宣布为获胜者（winner）。获胜神经元所表示的模型向量，以及位于获胜者周围指定半径中的那些与获胜者邻近的模型向量不断变化，使得它们能够更好地表示输入模式。使用所有的样本上对该过程不断地进行迭代，最终

完成的映射图成为对所有输入向量所占据空间的表示。

SOM 算法

对模型向量的拟合通过连续的回归过程而实现。如果 t 是样本的索引，由实向量描述的输入向量为 $\boldsymbol{x}(t)\in\mathbf{R}^n$，并且映射图中的每一个节点包含了一个模型向量 $\boldsymbol{m}(t)\in\mathbf{R}^n$，它与输入矢量大小相同。$\boldsymbol{m}(t)$ 的初始化可以是随机的或者以某种有序的方式进行，例如，沿着输入数据向量的两个主要特征向量所跨越的二维子空间。自组织算法可以用下面的步骤来解释，它通过竞争性的学习过程创建有序的映射图（假定一个 $N_1\times N_2$ 的网格）：

1. 随机或者有序地选择一个输入向量 $\boldsymbol{x}(t)$，然后对于 $k=1\cdots N_1\times N_2$，将它与所有的模型向量 $\boldsymbol{m}_k(t)$ 进行比较。识别出映射图上的最佳匹配节点，也就是说，模型向量在某种测度下（例如，欧氏距离或者 Pearson 相关系数）与输入向量最为相似的节点。这个最佳匹配节点 i 被宣布为获胜者。
2. 根据学习原理，获胜者的模型向量 $\boldsymbol{m}_i(t)$ 和它在目标网格中的一些相邻节点以逼近输入向量的方式被更新。SOM 学习过程的基本任务是：对于每个样本输入向量 $\boldsymbol{x}(t)$，在靶空间中对获胜者和与它相邻的节点进行调整，使得它们与源空间（输入数据空间）的 $\boldsymbol{x}(t)$ 接近。如果可以使用的输入样本数量有限，那么样本必须重复递交给 SOM 算法。学习过程中模型向量的适应可以根据下面的公式进行：

$$\boldsymbol{m}_k(t+1)=\boldsymbol{m}_k(t)+\alpha[\rho(i,k),t][\boldsymbol{x}(t)-\boldsymbol{m}_i(t)] \tag{5}$$

式中，t 是变量的离散化迭代索引；学习率 $\alpha(\rho,t)\in[0,1]$ 是一个定义了学习步长相对大小的标量因子，它依赖于迭代索引 t 和邻域函数 ρ。函数 ρ 指定了在映射网格中获胜节点 i 周围的邻域。在学习过程的起始阶段邻域半径可以选择得相当大，但是伴随着学习过程而逐步地缩小。这就保证了在开始时得到一个全局性的组织，而接近结束时，随着半径的变小，在映射图中对模型向量的局部修正将会更为精确。正是在这里提供了拓扑信息。邻近节点受到类似的更新，并因此结束对邻近输入模式的响应。邻域函数的一个例子为：

$$\rho(i,k)=e^{d_{i,k}^2/(2\sigma^2)} \tag{6}$$

式中，$d_{i,k}^2$ 是节点 i 和 k（网格）之间的距离；σ 是随着时间而逐渐下降的宽度参数。而且 $\alpha(\rho,i)$ 可以是，

$$\alpha(\rho,i)=\rho(i,k)(a/[1-bt])$$

式中，参数 a 控制步长；b 控制随着迭代索引 t 的下降速率。

基于 SOM 的聚类算法利用了从目标网格到源空间（输入数据空间）的映射。输入数据的各点被指派到最接近的（对于一个给定的距离定义）目标网格映射中，这样就得到了需要的聚类结果。

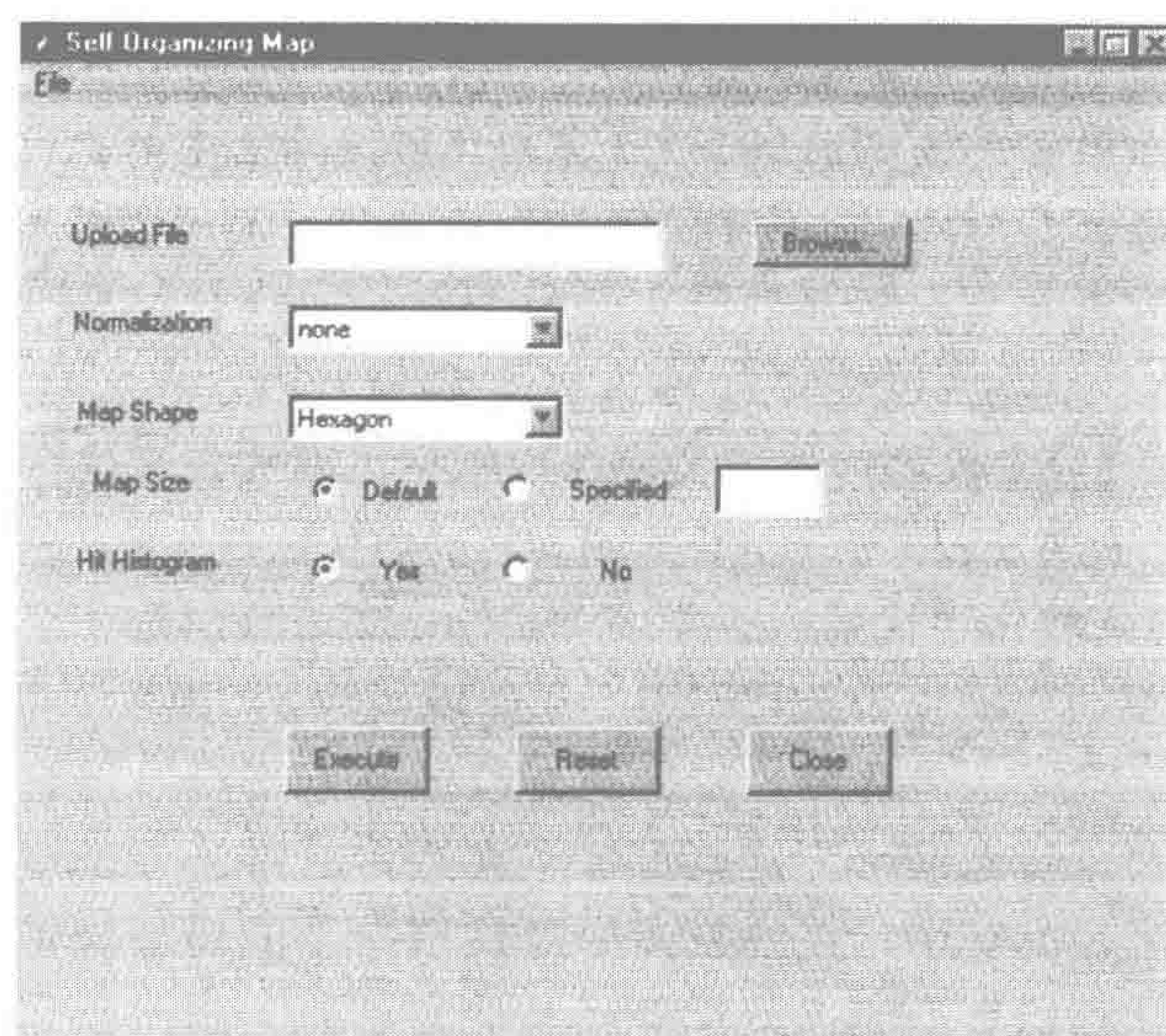
实现

许多基因表达谱实验使用 SOM 实现聚类功能（Tamayo et al. 1999；Toronen et al. 1999）。SOM 的一种 MATLAB 程序实现，最初由赫尔辛基科技大学计算机和信息科学实验室开发，经过修改后增加了运行程序基本功能的图形化用户界面（GUI）。这

里对该 SOM 程序进行介绍，使用了 Bittner 等（2000）讨论过的数据集。原始的 SOM 程序由赫尔辛基科技大学维护，包括全部文档都可以自由下载，在 http://www.cis.hut.fi/research/上可获取（Vesantao et al. 1999）。该程序已经在 Windows NT 4.0 平台上的 MATLAB v5.3 中测试过，高度建议系统具有快速的处理器和最低 64 Mb 的 RAM。因为 GUI 是在 Windows 平台下开发的，在其他环境下可能会有运行差异。安装软件包时，文件夹（名称为 SOM）将根据用户的选择移动到相应的位置，并且设置 MATLAB 的路径包含 SOM 文件夹。运行 MATLAB 环境，并键入

```
>>som _ launch
```

该命令产生一个 GUI，如图 7-49 所示。该 GUI 允许用户为他们的数据选择标准化方法（无、方差和自然对数）和映射网格的形状。作为选项，用户能够指定首选的映射图单元的数目，这将影响到最后映射图的最终大小。



588

图 7-49　SOM：图形化用户界面（GUI）

程序需要以 tab 分隔文本格式的输入，必须像图 7-50 所示的那样进行格式化，其中的行表示每个基因的表达数据，列表示每次实验中比值或者强度的测量值。第一行仅包含实验的数量或者输入空间维数，紧接着的一行以 #n 开始，包含了实验的标志。请注意基因的名称在最后一列中给出，没有与之相关联的列标志。

用户还可以选择使用一种命中直方图（hit histogram），它显示了在映射图中数据的分布，也可以省略它。SOM 产生的结果如图 7-51 和图 7-52 所示。图 7-51 显示了 U 矩阵可视化，它描述了邻近节点间的距离，因此代表了映射图中的聚类结构。

假定在 U 矩阵中的类别为具有低数值的一致性区域，U 矩阵中高数值表示了邻近的映射图单元之间有较大的距离，因此就构成了类别之间的边界。图 7-52 显示了包含在每个节点中数据点的平均表达谱。此外，为了使得用户能够识别每个节点中所包含的成分，全部基因名称列表和它们的位置（节点数字）一道输出到名为 som _ nodes.txt 的文件。

4

#n	UACC383	UACC457	UACC3093	UACC253
0.06	0.12	0.09	0.15	183950
0.06	0.05	0.05	0.1	486844
0.09	0.05	0.07	0.14	30966
0.25	0.24	0.24	0.25	244147
0.14	0.41	0.38	0.36	31687
0.31	0.63	0.05	0.84	813823
0.14	0.59	0.29	0.27	810603
0.46	0.23	0.5	0.71	179711
0.18	0.36	0.08	0.42	32609
0.2	0.5	0.22	0.36	114116
0.1	0.09	0.07	0.33	755663
0.15	0.15	0.18	0.23	357374
0.22	0.28	0.29	0.31	366830
0.51	0.12	0.21	0.24	71101
0.45	0.07	0.06	0.07	233721
0.09	0.06	0.06	0.13	760224

图 7-50　SOM 数据输入：Tab 分隔的文本

589

行表示了每个基因的表达数据，而列表示了每次实验中测量的比值或强度

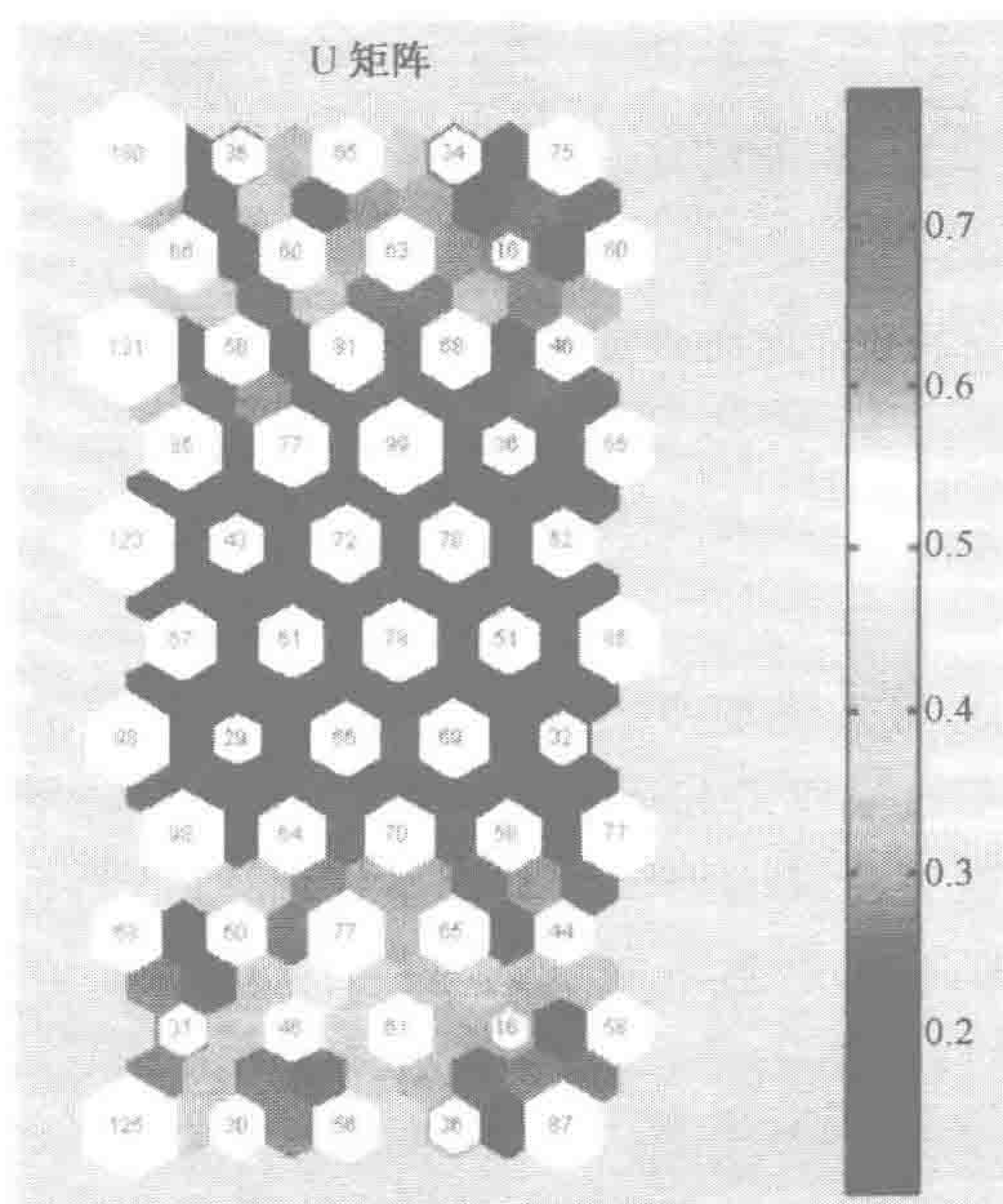


图 7-51　SOM 生成的结果：U 矩阵的可视化

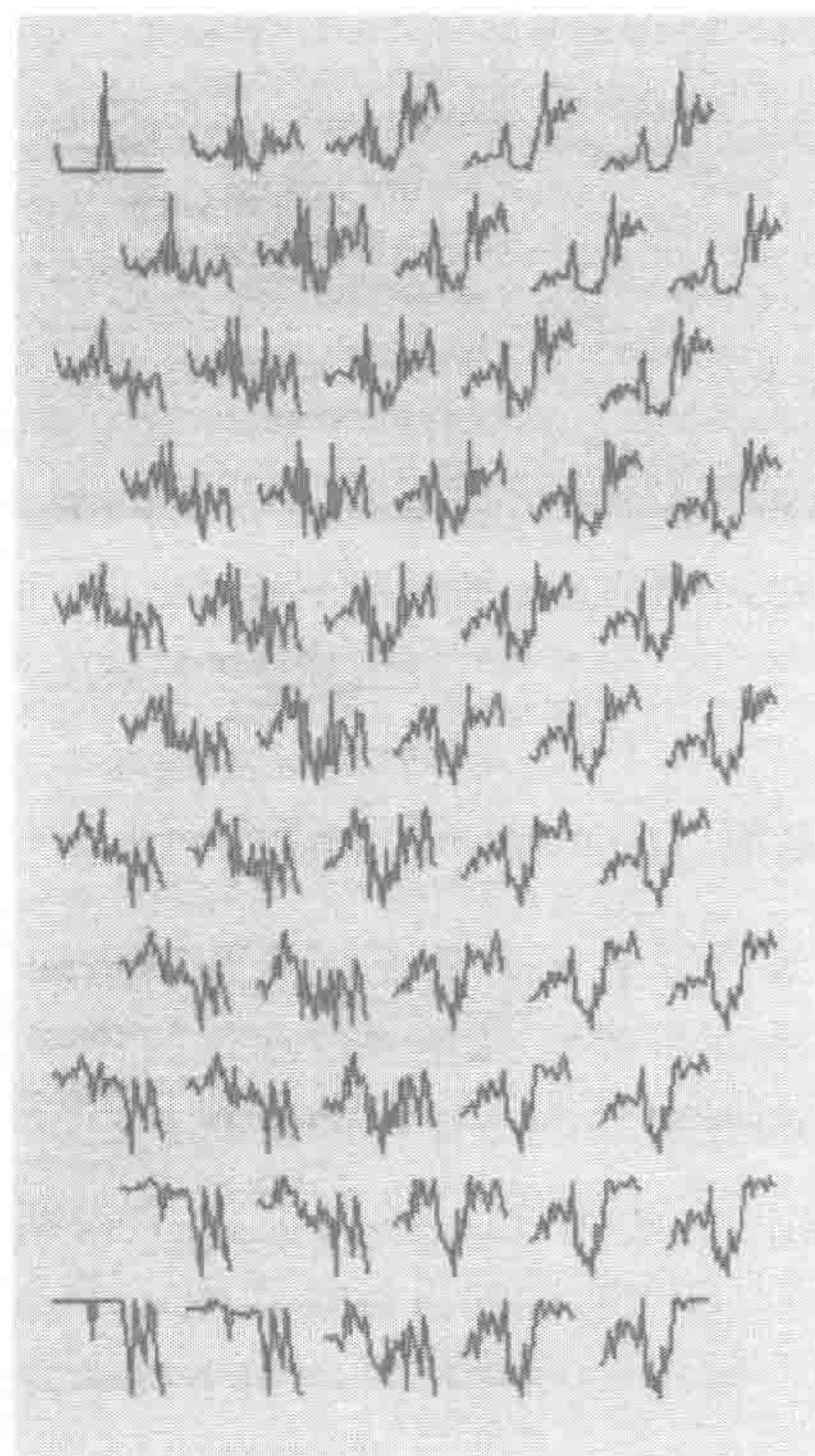


图 7-52　SOM 生成的结果：数据点的平均表达谱

只要用户选择，他们就能够通过包含在 som _ demo1-4中的 4 个例子了解 SOM 软件包的众多性能。对数据更为详尽的分析可以通过 MATLAB 命令行来进行，与之相伴的还有 SOM 工具包中所提供的丰富的数据分析函数（http://www. cis. hut. fi/projects/somtoolbox/documentation）。

590

信息栏：基因注释（gene annotation）

获得关于单个基因恰当的和容易理解的信息的能力对于微阵列实验的有效解释而言是至关重要的。在理想的情况下，每个基因都将有唯一的标识符，而且可以获得完整的分子信息，包含用一个经过注释的序列来描述启动子元件、内含子/外显子结构、剪切变体和相关的基因。通过与其他数据库的链接，从 PubMed 到蛋白质结构数据库，与该基因相关的生物学信息可以方便地得到。对构成微阵列的基因进行全面注释，在为共调控基因指定潜在功能时是非常有用的。这在利用可获得的已知基因的信息来分析尚未完全弄清楚的基因，特别是 EST 序列时，是尤其有用的（图 7-53）。

YOR124C	UBP2	protein degradation	ubiquitin-specific protease	cytoplasm
YBR082C	UBC4	protein degradation	ubiquitin conjugating enzyme	unknown
YPR158W	—	unknown	unknown	unknown
YGR142W	BTN2	unknown	unknown	unknown
YOL032W	—	unknown	unknown	unknown
YDR171W	HSP42	stress response	chaperone	cytoskeleton
YKR011C	—	unknown	unknown	unknown
YGL087C	MMS2	protein degradation	ubiquitin conjugating enzyme	unknown
YFR003C	—	unknown	unknown	unknown
YFR010W	UBP6	protein degradation	ubiquitin-specific protease	unknown
YKL213C	DOA1	protein degradation	unknown	unknown
YGR136W	—	unknown	unknown	unknown
YIL076W	SEC28	vesicle coat assembly	vesicle transporter	coatomer
YOR007C	SGT2	unknown	unknown	unknown
YNL007C	SIS1	protein folding	unknown	unknown
YOR027W	STI1	protein folding	unknown	unknown
YPL240C	HSP82	protein folding	chaperonin	unknown
YMR186W	HSC82	protein folding	chaperone	unknown
YER103W	SSA4	protein folding	chaperone	cytoplasm
YDR258C	HSP78	protein folding	chaperone	mitochondrial matrix
YLL026W	HSP104	protein folding	cochaperone	cytoplasm
YLR217W	—	unknown	unknown	unknown
YLR216C	CPR6	protein folding	peptidylprolyl isomerase	cytoplasm
YLR259C	HSP60	protein folding	chaperonin	mitochondrion
YDR214W	—	protein folding	unknown	unknown
YOR020C	HSP10	protein folding	chaperone	mitochondrial matrix
YNL281W	HCH1	protein folding	unknown	unknown
YJL021C	—	unknown	unknown	unknown
YAL028W	—	unknown	unknown	unknown
YJL034W	KAR2	protein folding	chaperone	endoplasmic reticulum lumen
YLL024C	SSA2	protein folding	chaperone	cytoplasm
YAL005C	SSA1	protein folding	chaperone	cytoplasm
YAL004W	—	unknown	unknown	unknown
YFL016C	MDJ1	protein folding	chaperone	mitochondrial inner membrane
YNL006W	LST8	transport	unknown	unknown

591

图 7-53 微阵列实验的一个输出结果，GO 描述字提供了帮助确定共调控基因潜在功能的信息（经许可复制自 Ashburner et al. 2000）（见彩图）

GO 组织（GO consortium）

综合来自于各种不同来源信息的一个基本部分是使用结构化的、受约束的词汇表。这样的词汇表可以允许自动化的检索，以便于过滤信息和得到关于在一次微阵列实验中获得的数据与已经存在的信息之间可能联系的有意义的推论。基因本体组织（gene ontology consortium）进行了一个尝试来解决其中一些问题（2001；Ashburner et al. 2000）。GO 项目的目标是产生一个全面的受约束术语集，可以用于描述所有物种中的基因。使用 MEDLINE 中用于科学文献描述的 MeSH 标题是这一概念中一个熟悉的例子（Delozier and Lingle 1992；Lowe and Barnett 1994）。GO 项目起步于开发用于模式物种数据库 FlyBase、小鼠基因组信息学（MGI）数据库和酵母（*Saccharomyces*）基因组数据库（SGD）的共享词汇表，使用术语来描述：

- **分子功能（molecular function）**：各个基因产物完成的基本任务，如钙活化的钾

离子通道（calcium-activated potassium channel）和 ATP 依赖的 DNA 解旋酶（ATP-dependent DNA helicase）。

- **生物学途径（biological process）**：通过分子功能的有序组合实现的主要生物学目标，如有丝分裂（mitosis）或者嘌呤代谢（purine metabolism）。
- **细胞内成分（cellular component）**：亚细胞结构、定位和大分子复合体，如细胞核（nucleus）、端粒（telomere）和起点识别复合体（origin recognition complex）。

GO 项目已经扩展到包括拟南芥（*Arabidopsis*，TAIR）、线虫（*Caenorhabditis elegans*，WormBase）和裂殖体酵母（*Schizosaccharomyces pombe*，PomBase）。另外，TIGR、Proteome、EBI、Compugen 以及 Celera 公司已经使用 GO 来注释细菌和人类基因。由 Proteome 所完成的对一些人类基因的注释可以通过 NCBI 上的 LocusLink 获得。

一系列的微阵列实验经常能够揭示具有相关表达谱的成组基因。根据来源于一组基因中已命名基因的功能注释信息，可以预测该组中具有很少、甚至还没有生物学信息存在的成员的可能功能。这在确定 EST 序列的潜在功能归属时尤其具有价值。此外，观察到一组共调控基因中多个成员具有共同的生物学效果，可以巩固特定的生物学过程正在为微阵列分析所揭示这一结论。

关于注释是如何提供有关 EST 序列的信息，并且揭示未预料到的复杂应答，描述血清对成纤维细胞刺激的时间过程实验提供了一个很好的例子（Iyer et al. 1999）。虽然这些细胞是在体外进行培养，反映复杂生物学应答的多组基因发生了表达，预示着刺激后的伤口愈合和组织重建反应。许多共调控的基因组中包含了参与组内已命名成员所属过程的 EST 序列。

GO 术语注释通常使用最可能的精确术语，但是存在一些场合，这时一个更为通用的术语能够提供关于一组基因所共有的更多信息。例如，如果与 DNA 损伤检查点（DNA damage checkpoint）、胞质分裂（cytokinesis）和 DNA 复制起始（DNA replication initiation）等术语有关的一组基因中的每一个均用有丝分裂细胞周期（mitotic cell cycle）的注释来代替，那么会更加易于解释。GO 组织正在计划创建和发布"GO Slim"（简化的基因本体术语），其中的注释将会转移到高层次的术语。更为灵活的调整 GO 注释特异性的方法对微阵列研究将非常有利，因此使用的高层次术语能够适合于即将进行的研究。

注释 cDNA 序列（annotating cDNA sequences）

借助于多种来源的信息，cDNA 克隆可以使用多种方式进行注释。迄今为止，最大的挑战是使得多种来源的信息吻合起来。大多数实验室通过 GenBank 存取号或者 IMAGE 克隆 ID 来识别 cDNA 克隆，但是这种做法并不标准。NCBI 提供了一些文件，用于说明来自不同项目中的 ID 之间是如何进行对应的（例如，见 LocusLink 的 ftp 站点 ftp://ncbi.nlm.nih.gov/refseq/LocusLink/）。

分析微阵列数据时，基因注释的更新会引起混淆。特别是在所有基因的精确定位尚未完全确定的基因组中，克隆和它们的基因之间对应关系每周都有可能产生新的表示。

尽可能采用提供最流行的和最新信息的注释系统进行工作是非常重要的。下面列举了一些相当流行的数据源：

dBEST（Database of Expressed Sequence Tags，表达序列标签数据库）：

http://www.ncbi.nlm.nih.gov/dbEST/index.html

dBEST 提供了来自上百种物种的 EST 的序列信息。

Swiss-Prot：http://www.expasy.ch/sprot/

Swiss-Prot 提供了受管理的蛋白质数据集。蛋白质注释的质量和细节水平很高，由它们产生的数据与其他数据库整合得很好。Swiss-Prot 的关键词本身就是方便的简短注释，能够提供关于蛋白质在细胞内功能的快捷信息。

GeneCards：http://bioinformatics.weizmann.ac.il/cards/ 或者

http://genome-www.stanford.edu/genecards/

GeneCards 是由魏茨曼科学研究院（Weizmann Institute of Science）建立的。它提供了关于人类基因及其在人类疾病中作用的非常简明的信息集。GeneCards 可以通过关键词或者被认可的基因名称进行检索。它以自动化的方式从多个外部数据源提取信息。

SOURCE：http://genome-www.stanford.edu/source

用于克隆和 EST 的斯坦福在线通用资源库（The Stanford Online Universal Resource for Clones and ESTs，SOURCE）综合了关于人类、小鼠和大鼠的信息。与 GeneCards 类似，它以自动化的方式从多个外部数据源提取信息。可以使用基因名称或克隆标识符进行检索，而且可以检索到关于克隆或者它们相关基因的信息。SOURCE 允许进行批量查询（同时使用多个标识符进行数据检索），这一点在分析一组共表达的基因时是非常方便的。

UniGene：http://www.ncbi.nlm.nih.gov/UniGene/

UniGene 将 GenBank 序列（包括 EST 序列）分割为非冗余的以基因为导向的聚类集合。UniGene 进行定期更新，这样做的优点是具有最新的信息，但是缺点是在克隆和聚类的类别之间有一些非常不稳定的关系。在每一个 UniGene 版本中，UniGene 的类别 ID 可以被重新进行分配。可以在 NCBI 的 LocusLink ftp 站点（ftp://ncbi.hlm.nih.gov/refseq/ LocusLink）找到关联 UniGene ID 与其他 ID 的表格。UniGene 当前为人类、大鼠、小鼠、母牛、斑马鱼、青蛙、小麦、水稻、大麦、玉米和水芹建立了聚类（表 7-10）。

593

表 7-10　基因注释的数据源

资　源	URL	信息类型
Unigene	www.ncbi.nlm.nih.gov/UniGene	将 EST 分组到唯一的、以基因为导向的各类别之中（人类、小鼠、大鼠、母牛和斑马鱼）
LocusLink(Pruit 和 Maglott 2001)	www.ncbi.nlm.nih.gov/LocusLink	关于基因受管理的序列和信息（人类、果蝇、小鼠、大鼠和斑马鱼）
dBEST(Boguski et al. 1993)	www.ncbi.nlm.nih.gov/dbEST	来自许多物种的 EST 序列
GeneMap99	www.ncbi.nlm.nih.gov/genemap	人类基因的作图信息
RHdb(Rodriguez-Tome 和 Lijnzaad 2001)	Corba.ebi.ac.uk/RHdb	人类、小鼠和大鼠的作图信息

续表

资　源	URL	信息类型
GeneCards(Rebhan et al. 1998)	Bioinformatics. weizmann. ac. il/cards	人类基因的功能和在疾病中的作用
SwissProt (Bairoch 和 Apweiler 2000)	www. expasy. ch/sprot	许多物种的注释的蛋白质序列
SOURCE	Genome-www. Stanford. edu/source	关于人类、小鼠和大鼠的基因和 cDNA 的汇编信息
小鼠基因组信息学(Mouse Genome Informatics, Blake et al. 2001)	www. informatics. jax/org	小鼠基因和序列的注释
FlyBase (FlyBase 组织 1999)	flybase. bio. indiana. edu	果蝇(*Drosophila melanogaster*)基因和序列的注释
TAIR(Huala et al. 2001)	www. arabidopsis. org	拟南芥(*Arabidopsis thaliana*)基因的生物学信息
WormBase(Stein et al. 2001)	www. wormbase. org	线虫(*Caenorhabditis elegans*)基因的生物学信息
酵母基因组数据库(*Saccharomyces* Genome Database, Ball et al. 2001)	genome-www. stanford. edu/Saccharomyces	酿酒酵母(*Saccharomyces cerevisiae*)的基因和序列的注释信息
基因本体组织(Gene Ontology Consortium, Ashburner et al. 2000)	www. geneontology. org	描述基因产物的过程、功能和定位的结构化的、受约束词汇表
Proteome	www. proteome. com	对酿酒酵母(*S. cerevisiae*)、粟酒裂殖酵母(*Schizosaccharomyces pombe*)和线虫(*C. elegans*)蛋白质的注释

信息栏：在研究人员之间共享微阵列数据

594

从事微阵列研究的实验室在其研究项目的早期常常会经历过重的数据负担（Mills and Gordon 2001；Mills et al. 2001）。随着微阵列技术更加容易使用和普及，这个问题只会变得更糟。如果世界上数以千计的实验室中产生的信息要进行传播，必须在存档格式和数据注释的标准方面达成一致。对于当前数据以非标准格式存储，常常位于分散的网站中作为补充信息，并且提供不同程度的实验信息这种现状，现在已有实质性的动机要对此加以改变（Becker 2001；Brazma et al. 2001）。

微阵列数据的发布非常重要，因为这样可以使得其他的研究人员能够对数据的质量和分析进行评价和验证，并允许超越论文中的方法而进行深入的数据挖掘。另外，通过提供大量可用于检验的数据，能够促进开发新的分析工具。与 DNA 测序项目产生的数据相比，微阵列实验的某些方面给结果共享造成了实质性的问题，这不仅在于对实验本身的描述，还有对点样在微阵列上序列的描述（Brazma et al. 2001）。从技术的角度来看，杂交（微阵列的制备和标记方法，其中包括扩增步骤和扫描平台）和使用的数据分析方法（如标准化方法）会极大地影响到结果和结论。来自竞争性杂交微阵列的数据是比较性的，而且测定的是相对表达水平而不是绝对的转录物数量。针对这个问题，需要

开发公用的表达标准或参考，以使得数据的普遍比较容易进行（见第 3 章，方案 12）。作为替代，对 RNA 丰度绝对测量而非基因表达比较性测量的关注日益增加，这将使得对特定细胞系和发育阶段的转录组有一个更加稳健的描述（Velculescu et al. 1997）。最后，在测量基因表达谱时，细胞或组织的状态都是重要的。对表达谱有主要影响的变化包括取样时机体的发育阶段或生理状态、组织取样的部位、细胞系所使用的生长条件等。

MGED

为了解决微阵列数据共享所引发的一些问题而成立了一个组织。MGED 小组是一个开放的讨论组，最初成立在英国剑桥举行的第一次微阵列基因表达数据库（Microarray Gene Expression Database）会议上（http://www.mged.org/）。这个小组的目标是推动关于 DNA 微阵列实验注释和数据表示标准的采纳，同时也要提出标准的实验对照和数据标准化方法。根本的目标是促进基因表达数据库的建立，建立不同来源的基因表达数据可比较性的标准，以及使得不同的基因表达数据库和数据分析软件能够协同工作。MGED 由来自许多学术性和私立研究所的研究人员和开发人员组成（例如，EBI、TIGR、Rosetta、Stanford 和 Harvard）。

MGED 中有两个平行但又互补的项目关注存储微阵列数据的内容、格式和交换。595 第一个是 MIAME（minimal information about a microarray experiment），该项目要定义关于一个微阵列实验最低限度的信息（Brazma et al. 2001）。第二个是 MAGEML（microArray gene expression markup language，微阵列基因表达标记语言），它的目标是定义一个用于报告微阵列信息的结构，可以用于数据的通讯和计算以及相关的注释。标记语言正在由一个大的合作组织进行定义并使用了 XML。

MIAME

MIAME1.0 是在 2001 年 3 月，在斯坦福大学召开 MGED 3 会议上被认可的（Kellam 2001）。MIAME 文件（http://www.mged.org/Annotations-wg/index.html）试图保存解释和重复结果所必需的全部信息（Brazma et al. 2001）。一个挑战是要指定保存了实验基本信息的必要信息集合，而不会给用户增添不必要的负担。尽管 MIAME 没有指定格式，一个关键的目标是将信息结构化使得能够进行有效的查询以及自动化数据分析和挖掘。这需要使用受约束的词汇表，而不是自由格式的文本。MIAME 标准描述了六类实验的注释：

- **实验设计：一组杂交的整体性描述**。MIAME 规范允许一组相关的微阵列结果归类在一起，并作为一个整体而进行描述。可以描述实验的类型（例如，时间序列或者疾病组织与健康组织之间的比较）和对阵列之间关系的详细描述。
- **微阵列设计**。除了关于点样在阵列上探针的序列信息，该类别的目的是记录微阵列平台和制备方案的详细情况。Part V 中讨论了这种类型信息的记录，涉及 LIM 系统和序列的建模。
- **样本**。对用于杂交样本的描述必须包括的主题有生物学样本的原始来源，取自何种生物，以及用于样本制备和标记的方案。

- **杂交**。MIAME 的杂交部分描述了杂交反应的标准技术参数，如缓冲液的浓度、温度和浓度。
- **检测**。来自对每次杂交扫描的数据必须包括生成这些结果所使用的硬件和软件，以及一组完整的结果和关于实验重复的信息。
- **标准化和对照**。必须描述用于数据标准化的策略和算法。此外，这部分中还记录了阵列上任何对照样点的位置信息。

MGED 小组致力解决的最困难问题之一是原始图像数据是否必须存储？储存原始图像文件保留了最大量的信息，使得后续分析中可以应用不同的标准化、图像提取和质量度量方法。正如 Geschwind（2001）最近的综述和表 7-11 中所作的总结那样，对来自微阵列实验输出结果的储存有许多种选择。在过渡时期建议储存原始数据文件；然而，这些文件非常大，会对于长期储存能力有着实质性的影响。

596

表 7-11　共享模式的优点和缺点

	原始图像	原始样点和背景数据	标准化的，处理过的数据（比值）	基因列表
优点				
整体上的灵活性	+++	++	+	
最小的偏差	+++	++	+	
允许重新分析	+++	++	+	
分析方法的比较	+++	++	++	
图像分割	+++			
图像分析	+++			
统计学上的灵活性	+++	+++	++	
数据挖掘的灵活性	+++	+++	++	
使用的方便性（分析需求最少）		+	++	+++
缺点				
需要大量存储	+++	+		
需要图像分析时间和软件	+++			
需要数据分析时间	+++	++	+	
潜在的平台之间不兼容性	+++	+		
标准化造成的偏差			+++	+++
一些统计学方法不能使用			++	+++
数据重新分析用处不大			+/−	+++
数据集合并用处不大			+	++
不是定量的				+++

经许可复制自 Geschwind [2001(© Macmillan Magazines Ltd.)]。

直到现在，还没有整合了 MIAME 的软件，虽然这种情况可能会迅速改变，因为 MIAME 或者它的变体，正在逐渐为研究界所接受。由于对微阵列数据的结构化记录正

在成为规范，毫无疑问，这些信息的提交将成为杂志和基金机构的必然要求。

用于存储和检索的公共数据库

当前正在开发一些用于微阵列数据的大规模公共存储和查询的数据库，包括 Gene Expression Omnibus（GEO，http://www.ncbi.nlm.nih.gov/geo/），ArrayExpress（http://www.ebi.ac.uk/arrayexpress），基因表达数据库（Gene Expression Database，GXD，http://www.informatics.jax.org/mgihome/GXD/aboutGXD.shtml）和 GeneX（http://www.ncgr.org/genex）（Brazma and Vilo 2000；Brazma et al. 2000，Gardiner-Garden and Littlejohn 2001）。设想 GEO 和 ArrayExpress 将用作存储不同来源的信息。基于 Jackson 实验室的 GXD，将集中于来自小鼠（*Mus musculus*）的表达数据。在本书写作时，所有的数据库都处在开发和接受数据的早期阶段。在 Gardiner-Garden 和 Littlejohn 的综述（2001）中评述了这些数据库的特性。

597

参考文献

Adams R. and Bischof L. 1994. Seeded region growing. *IEEE Trans. Pattern Anal. Machine Intell.* **16:** 641–647.

Alizadeh A.A., Ross D.T., Perou C.M., and van de Rijn M. 2001. Towards a novel classification of human malignancies based on gene expression patterns. *J. Pathol.* **195:** 41–52.

Alizadeh A.A., Eisen M.B., Davis R.E., Ma C., Lossos I.S., Rosenwald A., Boldrick J.C., Sabet H., Tran T., Yu X., Powell J.I., Yang L., Marti G.E., Moore T., Hudson Jr., J., Lu L., Lewis D.B., Tibshirani R., Sherlock G., Chan W.C., Greiner T.C., Weisenburger D.D., Armitage J.O., Warnke R., Staudt L.M., et al. 2000. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* **403:** 503–511.

Alon U., Barkai N., Notterman D.A., Gish K., Ybarra S., Mack D., and Levine A.J. 1999. Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. *Proc. Natl. Acad. Sci.* **96:** 6745–6750.

Alter O., Brown P.O., and Botstein D. 2000. Singular value decomposition for genome-wide expression data processing and modeling. *Proc. Natl. Acad. Sci.* **97:** 10101–10106.

Ashburner M., Ball C.A., Blake J.A., Botstein D., Butler H., Cherry J.M., Davis A.P., Dolinski K., Dwight S.S., Eppig J.T., Harris M.A., Hill D.P., Issel-Tarver L., Kasarskis A., Lewis S., Matese J.C., Richardson J.E., Ringwald M., Rubin G.M., and Sherlock G. 2000. Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium, *Nat. Genet.* **25:** 25–29.

Bairoch A. and Apweiler R. 2000. The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res.* **28:** 45–48.

Ball C.A., Jin H., Sherlock G., Weng S., Matese J.C., Andrada R., Binkley G., Dolinski K., Dwight S.S., Harris M.A., Issel-Tarver L., Schroeder M., Botstein D., and Cherry J.M. 2001. *Saccharomyces* Genome Database provides tools to survey gene expression and functional analysis data. *Nucleic Acids Res.* **29:** 80–81.

Becker K.G. 2001. The sharing of cDNA microarray data. *Nat. Rev. Neurosci.* **2:** 438–440.

Beucher S. and Meyer F. 1993. The morphological approach to segmentation: The watershed transformation. In *Mathematical morphology in image processing: Optical engineering* (ed. E. Dougherty), vol. 34, pp. 433–481. Marcel Dekker, New York.

Bittner M., Meltzer P., Chen Y., Jiang Y., Sefton E., Hendrix M., Radmacher M., Simon R., Yakhini Z., Ben-Dor A., Sampas N., Dougherty E., Wang E., Marincola F., Gooden C., Leuders J., Gatfelter A., Pollock P., Carpten J., Gillanders E., Leja D., Dietrich K., Beaudry C., Berens M., Alberts D., and Sondak V. 2000. Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature* **406:** 536–540.

Blake J.A., Eppig J.T., Richardson J.E., Bult C.J., and Kadin J.A. 2001. The Mouse Genome Database (MGD): Integration nexus for the laboratory mouse. *Nucleic Acids Res.* **29:** 91–94.

Boguski M.S., Lowe T.M., and Tolstoshev C.M. 1993. dbEST-database for "expressed sequence tags." *Nat. Genet.* **4:** 332–333.

Borg I. and Groenen P. 1997. *Modern multidimensional scaling: Theory and applications.* Springer, New York.

Box G.E.P., Hunter W.G., and Hunter J.S. 1978. *Statistics for experimenters: An introduction to design, data analysis, and model building.* Wiley, New York.

Brazma A. and Vilo J. 2000. Gene expression data analysis. *FEBS Lett.* **480:** 17–24.

Brazma A., Robinson A., Cameron G., and Ashburner M. 2000. One-stop shop for microarray data. *Nature* **403:** 699–700.

Brazma A., Hingamp P., Quackenbush J., Sherlock G., Spellman P., Stoeckert C., Aach J., Ansorge W., Ball C.A., Causton H.C., Gaasterland T., Glenisson P., Holstege F.C., Kim I.F., Markowitz V., Matese J.C., Parkinson H., Robinson A., Sarkans U., Schulze-Kremer S., Stewart J., Taylor R., Vilo J., and Vingron M. 2001. Minimum information about a microarray experiment (MIAME)—Toward standards for microarray data. *Nat. Genet.* **29:** 365–371.

Brown M.P., Grundy W.N., Lin D., Cristianini N., Sugnet C.W., Furey T.S., Ares Jr., M., and Haussler D. 2000. Knowledge-based analysis of microarray gene expression data by using support vector machines, *Proc. Natl. Acad. Sci.* **97:** 262–267.

Buhler J., Ideker T., and Haynor D. 2000. Dapple: Improved techniques for finding spots on DNA microarrays. University of Washington, Tech. Rep. UW-CSE-00-08-05 (http://bio.cs.washington.edu/online.html).

Callow M.J., Dudoit S., Gong E.L., Speed T.P., and Rubin E.M. 2000. Microarray expression profiling identifies genes with altered expression in HDL-deficient mice. *Genome Res.* **10:** 2022–2029.

Chen Y., Dougherty E.R., and Bittner M.L. 1997. Ratio-based decisions and the quantitative analysis of cDNA microarray images. *J. Biomed. Optics* **2:** 364–374.

Cleveland W.S. 1979. Robust locally weighted regression and smoothing scatterplots. *J. Am. Stat. Assoc.* **74:** 829–836.

Cox T.F. and Cox M.A.A. 2000. *Multidimensional scaling*, 2nd edition. CRC Press, Boca Raton, Florida.

De Gregorio E., Spellman P.T., Rubin G.M., and Lemaitre B. 2001. Genome-wide analysis of the *Drosophila* immune response by using oligonucleotide microarrays. *Proc. Natl. Acad. Sci.* **98:** 12590–12595.

Delozier E.P. and Lingle V.A. 1992. MEDLINE and MeSH: Challenges for end users. *Med. Ref. Serv. Q.* **11:** 29–46.

DeRisi J.L., Iyer V.R., and Brown P.O. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278:** 680–686.

Dhanasekaran S.M., Barrette T.R., Ghosh D., Shah R., Varambally S., Kurachi K., Pienta K.J., Rubin M.A., and Chinnaiyan A.M. 2001. Delineation of prognostic biomarkers in prostate cancer. *Nature* **412:** 822–826.

Dougherty E.R. 2001. Small sample issues for microarray-based classification. *Comp. Funct. Genomics* **2:** 28–34.

Dudoit S., Yang Y.H., Callow M.J., and Speed T.P. 2002. Statistical methods for identifying genes with differential expression in replicated cDNA microarray experiments. *Statistical Sinica* **12:** 111–139.

Eisen M.B. and Brown P.O. 1999. DNA arrays for analysis of gene expression. *Methods Enzymol.* **303:** 179–205.

Eisen M.B., Spellman P.T., Brown P.O., and Botstein D. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci.* **95:** 14863–14868.

Everitt B.S. and Dunn G. 1992. *Applied multivariate data analysis*. Oxford University Press, New York.

Finkelstein D.B., Gollub J., Ewing R., Sterky F., Somerville S., and Cherry J.M. 2000. Iterative linear regression by sector: Renormalization of cDNA microarray data and cluster analysis weighted by cross homology. CAMDA'00 (Critical Assessment of Microarray Data Analysis Techniques), Duke Bioinformatics Shared Resource (http://bioinformatics.duke.edu/CAMDA/CAMDA00/Abstracts).

Fisher R.A. 1926. The arrangement of field experiments. *J. Min. Agric. Gr. Br.* **33:** 505–513.

FlyBase Consortium. 1999. The FlyBase database of the *Drosophila* genome projects and community literature. *Nucleic Acids Res.* **27:** 85–88.

Freedman D., Pisani R., and Purves R. 1998. *Statistics*, 3rd edition. Norton, New York.

Gardiner-Garden M. and Littlejohn T.G. 2001. A comparison of microarray databases. *Brief Bioinform.* **2:** 143–158.

Gasch A.P., Spellman P.T., Kao C.M., Carmel-Harel O., Eisen M.B., Storz G., Botstein D., and Brown P.O. 2000. Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell* **11:** 4241–4257.

The Gene Ontology Consortium. 2001. Creating the gene ontology resource: Design and implementation. *Genome Res.* **11:** 1425–1433.

Geschwind D.H. 2001. Sharing gene expression data: An array of options. *Nat. Rev. Neurosci.* **2:** 435–438.

Glonek G.F.V. and Solomon P.J. 2002. Factorial designs for microarray experiments. Technical Report, Department of Applied Mathematics, University of Adelaide.

Golub T.R., Slonim D.K., Tamayo P., Huard C., Gaasenbeek M., Mesirov J.P., Coller H., Loh M.L., Downing

599 J.R., Caligiuri M.A., et al. 1999. Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring, *Science* **286:** 531–537.

Green P.E. and Rao V.R. 1972. *Applied multidimensional scaling.* Dryden Press, Hinsdale, Illinois.

Green P.E., Carmone F.J., and Smith S.M. 1989. *Multidimensional scaling: Concepts and applications.* Allyn and Bacon, Needham Heights, Massachusetts.

Hedenfalk I., Duggan D., Chen Y., Radmacher M., Bittner M., Simon R., Meltzer P., Gusterson B., Esteller M., Kallioniemi O.P., Wilfond B., Borg A., and Trent J. 2001. Gene-expression profiles in hereditary breast cancer. *N. Engl. J. Med.* **344:** 539–548.

Heyer L.J., Kruglyak S., and Yooseph S. 1999 Exploring expression data: Identification and analysis of coexpressed genes. *Genome Res.* **9:** 1106–1115.

Huala E., Dickerman A.W., Garcia-Hernandez M., Weems D., Reiser L., LaFond F., Hanley D., Kiphart D., Zhuang M., Huang W., Mueller L.A., Bhattacharyya D., Bhaya D., Sobral B.W., Beavis W., Meinke D.W., Town C.D., Somerville C., and Rhee S.Y. 2001. The *Arabidopsis* Information Resource (TAIR): A comprehensive database and web-based information retrieval, analysis, and visualization system for a model plant. *Nucleic Acids Res.* **29:** 102–105.

Hughes J.D., Estep P.W., Tavazoie S., and Church G.M. 2000. Computational identification of *cis*-regulatory elements associated with groups of functionally related genes in *Saccharomyces cerevisiae. J. Mol. Biol.* **296:** 1205–1214.

Ihaka R. and Gentleman R. 1996. R: A language for data analysis and graphics. *J. Comput. Graph. Stat.* **5:** 299–314.

Iyer V.R., Eisen M.B., Ross D.T., Schuler G., Moore T., Lee J.C., Trent J.M., Staudt L.M., Hudson Jr., J., Boguski M.S., Lashkari D., Shalon D., Botstein D., and Brown P.O. 1999. The transcriptional program in the response of human fibroblasts to serum. *Science* **283:** 83–87.

Kellam P. 2001. Microarray gene expression database: Progress towards an international repository of gene expression data. *Genome Biol.* **2:** 4011.1–4011.3.

Kepler T.B., Crosby L., and Morgan K.T. 2000. Normalization and analysis of DNA microarray data by self-consistency and local regression. Santa Fe Institute (http://www.santafe.edu/sfi/publications/Abstracts/00-09-55abs.html).

Kerr M.K. and Churchill G.A. 2001a. Experimental design for gene expression microarrays. *Biostatistics* **2:** 183–201.

———. 2001b. Statistical design and the analysis of gene expression microarray data. *Genet. Res.* **77:** 123–128.

Kerr M.K., Martin M., and Churchill G.A. 2000. Analysis of variance for gene expression microarray data. *J. Computat. Biol.* **7:** 819–837.

Khan J., Simon R., Bittner M., Chen Y., Leighton S.B., Pohida T., Smith P.D., Jiang Y., Gooden G.C., Trent J.M., and Meltzer P.S. 1998. Gene expression profiling of alveolar rhabdomyosarcoma with cDNA microarrays. *Cancer Res.* **58:** 5009–5013.

Khan J., Wei J.S., Ringner M., Saal L.H., Ladanyi M., Westermann F., Berthold F., Schwab M., Antonescu C.R., Peterson C., and Meltzer P.S. 2001. Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. *Nat. Med.* **7:** 673–679.

Kihara C., Tsunoda T., Tanaka T., Yamana H., Furukawa Y., Ono K., Kitahara O., Zembutsu H., Yanagawa R., Hirata K., Takagi T., and Nakamura Y. 2001. Prediction of sensitivity of esophageal tumors to adjuvant chemotherapy by cDNA microarray analysis of gene-expression profiles. *Cancer Res.* **61:** 6474–6479.

Kohonen T. 1995. *Self-organizing maps.* Springer, Berlin, Heidelberg.

Lashkari D.A., DeRisi J.L., McCusker J.H., Namath A.F., Gentile C., Hwang S.Y., Brown P.O., and Davis R.W. 1997. Yeast microarrays for genome wide parallel genetic and gene expression analysis. *Proc. Natl. Acad. Sci.* **94:** 13057–13062.

Lee M.L., Kuo F.C., Whitmore G.A., and Sklar J. 2000. Importance of replication in microarray gene expression studies: Statistical methods and evidence from repetitive cDNA hybridizations. *Proc. Natl. Acad. Sci.* **97:** 9834–9839.

Lockhart D.J., Dong H., Byrne M.C., Follettie M.T., Gallo M.V., Chee M.S., Mittmann M., Wang C., Kobayashi M., Horton H., and Brown E.L. 1996. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Biotechnol.* **14:** 1675–1680.

Lowe H.J. and Barnett G.O. 1994. Understanding and using the medical subject headings (MeSH) vocabulary to perform literature searches. *J. Am. Med. Assoc.* **271:** 1103–1108.

Mills J.C. and Gordon J.I. 2001. A new approach for filtering noise from high-density oligonucleotide microarray datasets. *Nucleic Acids Res.* **29:** E72.

Mills J.C., Roth K.A., Cagan R.L., and Gordon J.I. 2001. DNA microarrays and beyond: Completing the journey from tissue to cell. *Nat. Cell Biol.* **3:** E175–E178.

Oja E. and Kaski S., eds. 1999. *Kohonen maps*. Elsevier, The Netherlands.

Perou C.M., Sorlie T., Eisen M.B., van de Rijn M., Jeffrey S.S., Rees C.A., Pollack J.R., Ross D.T., Johnsen H., Akslen L.A., Fluge O., Pergamenschikov A., Williams C., Zhu S.X., Lonning P.E., Borresen-Dale A.L., Brown P.O., and Botstein D. 2000. Molecular portraits of human breast tumours. *Nature* **406:** 747–752.

Pomeroy S.L., Tamayo P., Gaasenbeek M., Sturla L.M., Angelo M., McLaughlin M.E., Kim J.Y., Goumnerova L.C., Black P.M., Lau C., Allen J.C., Zagzag D., Olson J.M., Curran T., Wetmore C., Biegel J.A., Poggio T., Mukherjee S., Rifkin R., Califano A., Stolovitzky G., Louis D.N., Mesirov J.P., Lander E.S., and Golub T.R. 2002. Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature* **415:** 436–442.

Pruitt K.D. and Maglott D.R. 2001. RefSeq and LocusLink: NCBI gene-centered resources. *Nucleic Acids Res.* **29:** 137–140.

Rebhan M., Chalifa-Caspi V., Prilusky J., and Lancet D. 1998. GeneCards: A novel functional genomics compendium with automated data mining and query reformulation support. *Bioinformatics* **14:** 656–664.

Rodriguez-Tome P. and Lijnzaad P. 2001. RHdb: The Radiation Hybrid database. *Nucleic Acids Res.* **29:** 165–166.

Schadt E., Li C., Ellis B., and Wong W.H. 2001. Feature extraction and normalization algorithms for high-density oligonucleotide gene expression array data (see http://preprints.stat.ucla.edu/303/). *J. Cell. Biochem.* (suppl.) **37:** 120–125.

Schena M., Shalon D., Davis R.W., and Brown P.O. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray (comments) *Science* **270:** 467–470.

Schena M., Shalon D., Heller R., Chai A., Brown P.O., and Davis R.W. 1996. Parallel human genome analysis: Microarray-based expression monitoring of 1000 genes. *Proc. Natl. Acad. Sci.* **93:** 10614–10619.

Schiffman S.S., Reynolds M.L., and Young F.W. 1981. *Introduction to multidimensional scaling: Theory, method and applications*. Academic Press, New York.

Sherlock G., Hernandez-Boussard T., Kasarskis A., Binkley G., Matese J.C., Dwight S.S., Kaloper M., Weng S., Jin H., Ball C.A., Eisen M.B., Spellman P.T., Brown P.O., Botstein D., and Cherry J.M. 2001. The Stanford Microarray Database. *Nucleic Acids Res.* **29:** 152–155.

Soille P. 1999. *Morphological image analysis: Principles and applications*. Springer, Berlin.

Spellman P.T., Sherlock G., Zhang M.Q., Iyer V.R., Anders K., Eisen M.B., Brown P.O., Botstein D., and Futcher B. 1998. Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. *Mol. Biol. Cell* **9:** 3273–3297.

Stein L., Sternberg P., Durbin R., Thierry-Mieg J., and Spieth J. 2001. WormBase: Network access to the genome and biology of *Caenorhabditis elegans*. *Nucleic Acids Res.* **29:** 826.

Tamayo P., Slonim D., Mesirov J., Zhu Q., Kitareewan S., Dmitrovsky E., Lander E.S., and Golub T.R. 1999. Interpreting patterns of gene expression with self-organizing maps: Methods and application to hematopoietic differentiation. *Proc. Natl. Acad. Sci.* **96:** 2907–2912.

Tavazoie S., Hughes J.D., Campbell M.J., Cho R.J., and Church G.M. 1999. Systematic determination of genetic network architecture. *Nat. Genet* **22:** 281–285.

Toronen P., Kolehmainen M., Wong G., and Castren E. 1999. Analysis of gene expression data using self-organizing maps. *FEBS Lett.* **451:** 142–146.

Troyanskaya O., Cantor M., Sherlock G., Brown P., Hastie T., Tibshirani R., Botstein D., and Altman R.B. 2001. Missing value estimation methods for DNA microarrays. *Bioinformatics* **17:** 520–525.

Tseng G.C., Oh M.K., Rohlin L., Liao J.C., and Wong W.H. 2001. Issues in cDNA microarray analysis: Quality filtering, channel normalization, models of variations and assessment of gene effects. *Nucleic Acids Res.* **29:** 2549–2557.

Tukey J.W. 1977. *Exploratory data analysis.* Addison-Wesley, Boston, Massachusetts.

Velculescu V.E., Zhang L., Zhou W., Vogelstein J., Basrai M.A., Bassett Jr., D.E., Hieter P., Vogelstein B., and Kinzler K.W. 1997. Characterization of the yeast transcriptome. *Cell* **88:** 243–251.

Vesanto J., Himberg J., Alhoniemi E., and Parhankangas J. 1999. Self-organizing map in Matlab: The SOM Toolbox, pp. 35–40. In *Proceedings of the Matlab DSP Conference*, November, Espoo, Finland (http://www.cis.hut.fi/projects/somtoolbox/documentation/).

Vincent L. and Soille P. 1991. Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *IEEE Trans. Pattern Anal. Machine Intell.* **13:** 583–598.

Wang X., Ghosh S., and Guo S.W. 2001. Quantitative quality control in microarray image processing and data acquisition. *Nucleic Acids Res.* **29:** E75.

Wodicka L., Dong H., Mittmann M., Ho M.H., and Lockhart D.J. 1997. Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nat. Biotechnol.* **15:** 1359–1367.

Wolfinger R.D., Gibson G., Wolfinger E.D., Bennett L., Hamadeh H., Bushel P., Afshari C., and Paules R.S. 2001. Assessing gene significance from cDNA microarray expression data via mixed models. *J. Comput. Biol.* **8:** 625–637.

Yang Y.H., Buckley M.J., and Speed T.P. 2001a. Analysis of cDNA microarray images. *Brief. Bioinformatics* **2:** 341–349.

Yang Y.H., Buckley M.J., Dudoit S., and Speed T.P. 2000. Comparison of methods for image analysis on cDNA microarray data. Department of Statistics, University of California Tech. Rep 584 (http://www.stat.berkeley.edu/tech-reports/index.html).

———. 2002a. Comparison of methods for image anlaysis of cDNA microarrays. *J. Computat. Graph. Stat.* **11:** 108–136.

Yang Y.H., Dudoit S., Luu P., and Speed T.P. 2001b. Normalization for cDNA microarray. Microarrays: Optical technologies and informatics. *SPIE* (The International Society for Optical Engineering) **4266:** 141–152.

Yang Y.H., Dudoit S., Luu P., Lin D., Peng V., Ngai J., and Speed T.P. 2002b. Normalization for cDNA microarray: A robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res.* **30:** e15.

Yates F. 1937. The design and analysis of factorial experiments. Imperial Bureau of Soil Science, Tech. commun. no. 35, Harpenden, England.

Young F.W. 1987. *Multidimensional scaling: History, theory and applications* (ed. R.M. Hamer). Lawrence Erlbaum Assoc., Hillsdale, New Jersey.

因特网资源

For additional WWW addresses, see also Table 7-4 in Part 2 (Image Analysis) and Table 7-10 in the information panel on **GENE ANNOTATION**

http://bioinformatics.weizmann.ac.il/cards/ Genecards(TM) database of human genes, their products and their involvement in diseases, Weizmann Institute of Science.

http://cm.bell-labs.com/cm/ms/departments/sia/wsc/smoothsoft.html Local Regression Software, Bell Labs Innovations, Lucent Technologies.

http://genex.sourceforge.net/ GeneX open source gene expression database.

http://genome-www.stanford.edu/genecards/ Genecards(TM) database of human genes, their products and their involvement in diseases.

http://genome-www.stanford.edu/microarray The Stanford Microarray Database.

http://genome-www.stanford.edu/source The Stanford Online Universal Resource for Clones and ESTs (SOURCE).

http://microarrays.org/software.html The microarrays.org software page.

http://pevsnerlab.kennedykrieger.org/snomad.htm SNOMAD: (Standardization and *NO*rmalization of *M*icro*A*rray *D*ata) program.

http://rana.lbl.gov Lawrence Berkeley National Laboratory, University of California at Berkeley, Michael Eisen lab home page.

http://rana.lbl.gov/EisenSoftware.htm ScanAlyze. Michael Eisen lab software.

http://rana.lbl.gov/manuals/ClusterTreeView.pdf Cluster and TreeView Manual written by Michael Eisen.

http://www.axon.com/GN_GenePixSoftware.html Axon Instruments, Inc GenePix Pro 4.0 software.

http://www.cis.hut.fi/projects/somtoolbox/documentation/ SOM Toolbox for Matlab.

http://www.cis.hut.fi/research/ The Helsinki University of Technology Laboratory of Computer and Information Science Neural Networks Research Centre.

http://www.cmis.csiro.au/iap/Spot/spotmanual.htm Buckley M.J. 2000. The Spot user's guide—Software package for the analysis of microarray images. CSIRO Mathematical and Information Sciences, Australia.

http://www.ebi.ac.uk/arrayexpress/ ArrayExpress public database for microarray based gene expression data.

http://www.expasy.ch/sprot/ SWISS-PROT curated protein sequence database.

http://www.fruitfly.org/expression/immunity Genome-wide gene expression of *Drosophila* in response to immune challenge, web supplement to De Gregorio et al. 2001.

http://www.informatics.jax.org/mgihome/GXD/aboutGXD.shtml The Gene Expression Database community resource for gene expression information from the laboratory mouse.

http://www.mathworks.com/products/matlab/ The Mathworks introduction to MATLAB 6.1.

http://www.mathworks.com/products/statistics/ .The Mathworks Statistics Toolbox 3.

http://www.mged.org/ The Microarray Gene Expression Data Group home page.

http://www.mged.org/Annotations-wg/index.html Home page for the microarray annotations working group.

http://www.ncbi.nlm.nih.gov/dbEST/index.html Expressed Sequence Tags database.

http://www.ncbi.nlm.nih.gov/geo/ The NCBI Gene Expression Omnibus gene expression and hybridization array data repository.

http://www.ncbi.nlm.nih.gov/UniGene/ The UniGene experimental system for automatically partitioning GenBank sequences into a nonredundant set of gene-oriented clusters.

http://www.ncgr.org/genex The National Center for Genome Resources (NCGR) GeneX(TM) home page.

http://www.r-project.org/ The R Project for Statistical Computing.

http://www2.stratagene.com/gc/ Stratagene's GeneConnection web portal.

第8章　组织微阵列

Juha Kononen[*]，**Galen Hostetter**[*]（[*]癌症遗传学研究部，国家人类基因组研究所，国立卫生院，贝塞斯达，马里兰州 20892-4770），**Guido Sauter**[†]，**Olli-P. Kallioniemi**[*]（[†]病理学研究所，巴塞尔大学，瑞士），**Karl Brand and Melanie Trivett Peter MacCallum**（癌症中心，东墨尔本，维多利亚 3002，澳大利亚）

603

导言

组织微阵列技术与其他微阵列技术的原理相同，也就是说，都是在一个小的区域（一张载玻片）上制备高密度的（100～1000个标本）微小阵列单元（这里特指组织）。组织微阵列（tissue microarray，TMA）的制备是从存档的组织标本中得到小的圆柱形样品（中心组织检查），然后将数百个芯样以阵列形式高密度的包埋于载样蜡块中（图8-1）。一个典型的TMA可以包含数百个芯样，分别来自特定类型的肿瘤、特定的组织，或者是胚胎发育的特定时期。更专业的TMA可以用反映一种肿瘤在不同发生时期的病理标本来制备，或者是包含了特定临床信息的标本（如预后或是对治疗的反应），或者是原发肿瘤和转移灶的成对标本。

组织微阵列技术的主要优点是在一张TMA切片上进行单次实验可以得到多达1000个供体标本的分子特性。与此相比，在常规的分子病理学检查中，每张切片通常只包含一种标本的单张切片。应用常规方法，如分析500个肿瘤，就需要对500张切片染色。

每个TMA蜡块大约可以进行200次连续切片。对每一供体标本重复取样，并将所得到的芯样排列在一系列TMA蜡块中相同的x-y坐标位置，就可以复制大量的TMA。以预先确定的方式精确排列芯样，有助于发展适用于供体标本阵列的自动图像分析方

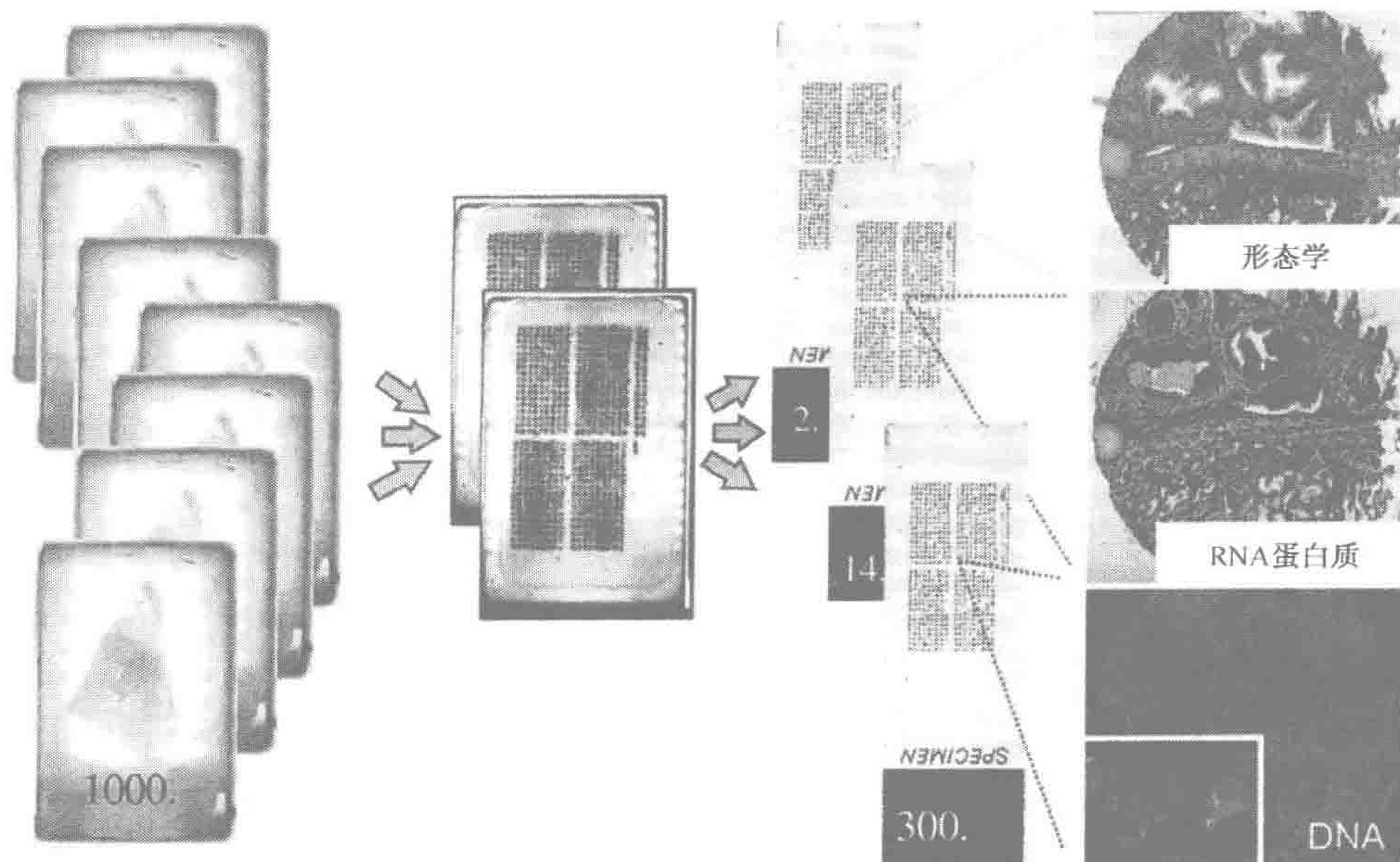


图 8-1 TMA 制备和后续分析演示

经过甲醛固定和石蜡包埋的组织芯样被转移到 TMA 蜡块。多个 TMA 蜡块可以同时制备，并且每个 TMA 蜡块可以连续切片约 200 次。最终得到的全部切片在相同的坐标位置上有着同样的组织。每张切片均可用于多种不同类型的分子检测，如用苏木精或者是曙红染色观察组织的形态学、mRNA 原位杂交（ISH）、蛋白质的免疫组化（IHC），以及用荧光原位杂交（FISH）分析遗传改变。（经许可复制自 Kallioniemi et al. 2001，© 牛津大学出版社）

法。复制可以显著提高应用每套标本进行实验的数量。由五个相同的 TMA 蜡块可以得到超过 1000 片相同的切片。

用于制备 TMA 的芯样要能代表原有病理标本的细胞组分和结构。如果不具备足够的代表性，那么与原始标本分析相比，TMA 可能会低估一些指标，比如基因表达中某种特定改变的频率。加大每个标本的芯样尺寸和数量会改善这种代表性，从而增加 TMA 的统计能力。然而，病理学标本这种珍贵资源必须仔细保存和经济使用。因此来自新鲜或者存档组织的芯样在数量和尺寸上必须尽可能的少和小。这些组织作为病理学档案保存，研究人员可以使用它们。许多研究机构要求研究人员在出于研究目的使用这些病理学材料之前，必须得到伦理学上的许可。

用于芯样活检的针可以购买到，直径通常为 0.6、1.0、1.5、2.0 mm 和 3.0 mm。除非组织的均一度非常高，否则直径小于 0.6 mm 的芯样没有足够的代表性；而直径大于 3 mm 的芯样不利于阵列的制备。应用 0.6 mm 的芯样对原始标本的浪费程度可以降至最低，因此在保留了形态学信息的同时也保护了独特而珍贵的研究资源。显而易见，目标应该是通过最少数量的最小尺寸芯样，得到与原始样品相比较可以接受的一致性。这样，病理学上均匀的组织，如肝脏，通过单个 0.6 mm 芯样活检，就可以获得精确的代表性。与此相比，对包含了许多不同类型细胞和结构的非均匀组织而言，多个 0.6

mm 芯样比单个 2 mm 芯样的代表性要好（Camp et al. 2000；Gillett et al. 2000；Sallinen et al. 2000；Hoos et al. 2001；Nocito et al. 2001；Rimm et al. 2001）。对于中等复杂程度的组织，我们推荐在样品上相距较远的部位选取两到三个芯样。

组织阵列制备过程中，芯样的获取以及后续的排列，都可以使用仪器或者是手工完成。应用精密仪器制备阵列改善了阵列的质量，并使得阵列制备和图像分析自动化成为可能。许多公司，如 Beecher Instruments（Silver Spring，Maryland）和 Chemicon International（Termecula，California）出售用于制备 TMA 的精密仪器。对芯样数目少于150个的阵列，用价廉的生物穿刺活检仪器直接制备，使得任何标准的病理实验室能采用这一技术。

TMA 代表了唯一可以高通量研究经过甲醛固定的存档组织材料的技术，它通常用于验证常规 cDNA 微阵列，或者其他基因组和蛋白质组筛选方法鉴定的候选基因的表达谱（图 8-2）（Bubendorf et al. 1999；Barlund et al. 2000；Moch et al. 2001）。用于通常“整块”病理标本的所有研究技术可以几乎不做修改而应用于 TMA。可用于筛选 TMA 来寻找特定分子改变的技术包括分子和组织学分析，如 DNA 靶序列的荧光原位杂交（FISH）、RNA 原位杂交（ISH）（见方案 2）、聚合酶链反应（PCR）原位杂交、免疫细胞化学和免疫组织化学（见方案 3）。尽管通常存档的甲醛固定组织样品中核酸性质可能会发生改变，这些组织样品代表了一个从存档组织块中获取临床和后续信息的强有力研究资源。研究显示，经过 10%缓冲甲醛固定的存档组织块，在 20～50 年后仍旧保持了免疫反应活性。因此我们推荐使用 10%缓冲甲醛，这样可以在长期的储藏后保持免疫反应活性。取样和排列阵列过程只是制备 TMA 所需工作的一部分。大量的精力花费在组织挑选和处理那些用于阵列的原始组织块（方案 1）。

606

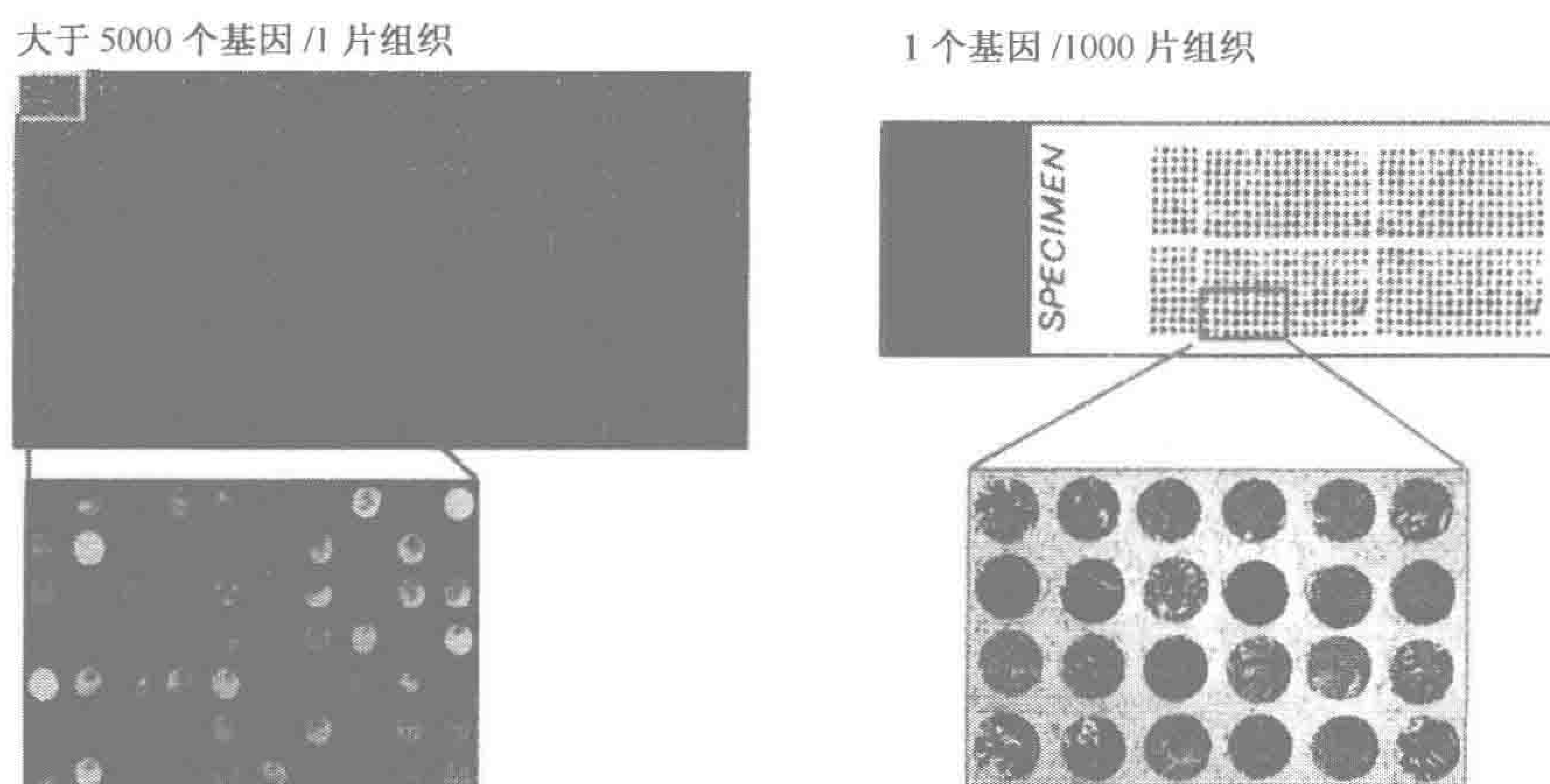


图 8-2 组织微阵列是对从各种基因组检查中，如 cDNA 微阵列得到的发现进行验证，区分优先次序和扩展的有力工具

每次 cDNA 微阵列试验可以分析一个样品中最多 5000 到 50 000 个基因。与此相比，TMA 每次可以分析一个靶标，但是每张载玻片上最多却可以容纳 1000 片组织切片。（经许可复制自 Kallioniemi 等，2001，牛津大学出版社）

用一种或数种这些技术，从少量新鲜采集样品分析中得到的结果，如表达谱分析，可以迅速扩展到大量的存档样品，并且同时可以根据组织结构的关系得到确认(Kononen et al. 1998；Bubendorf et al. 1999；Hedenfalk et al. 2001)。来自不同标记物的数据可以综合在一起来评价多参数分子谱图的能力，这可能是基于数十或者上百个参数。越来越多的研究报道证实了这一方法，并证明了 TMA 技术在分子病理学和转录肿瘤研究中的应用价值。

方案 1

制备组织微阵列

Juha Kononen[*]，**Galen Hostetter**[*]，**Guido Sauter**[†]，**Olli-P Kallioniemi**[*]（[*]癌症遗传学研究部，国家人类基因组研究所，国立卫生院，贝塞斯达，马里兰州 20892-4770；[†]病理学研究所，巴塞尔大学，瑞士）

本方案介绍了组织微阵列（TMA）的制备，并提供了疑难解答的推荐方法。该过程的流程图见图 8-3。

607

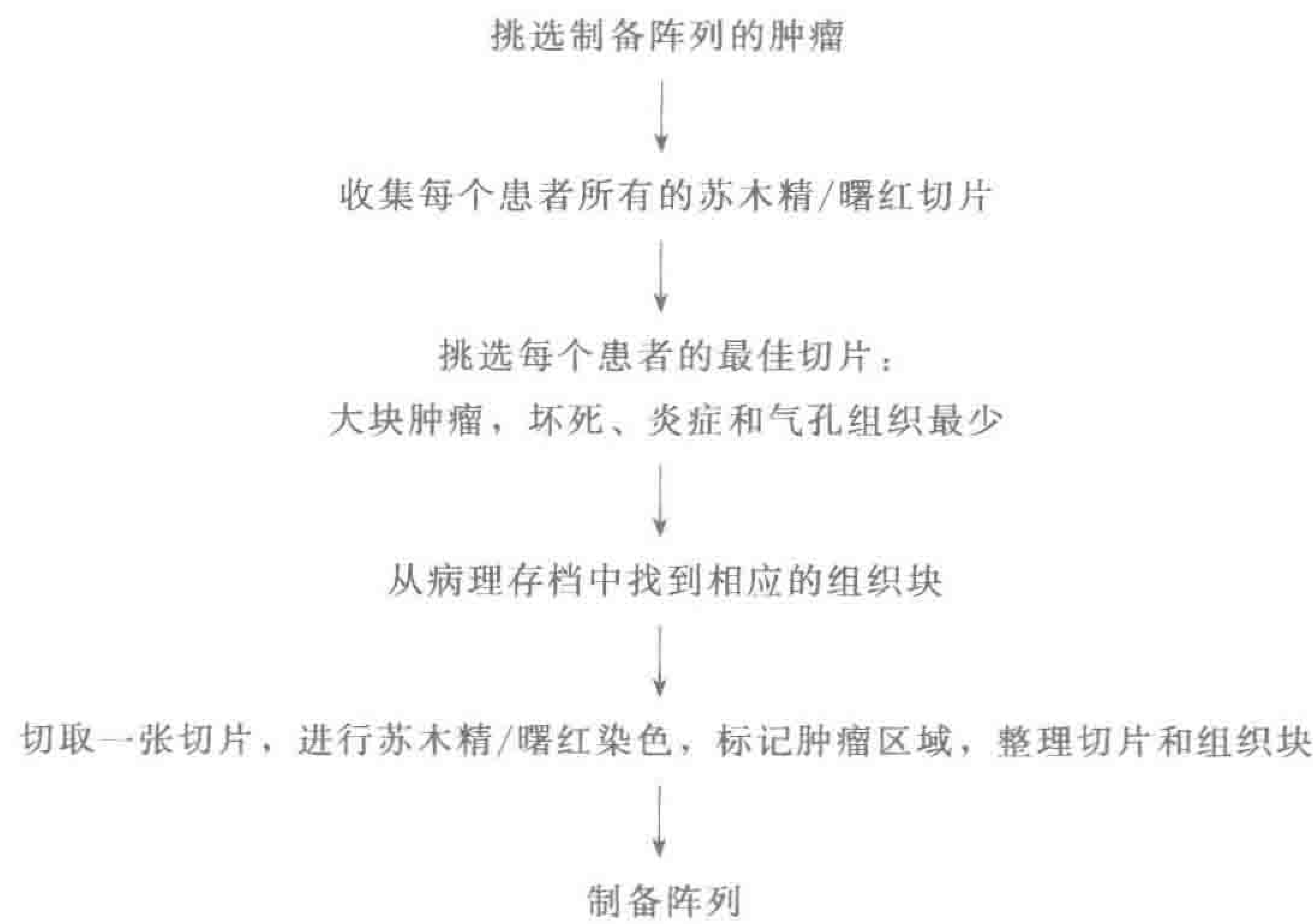


图 8-3　制备组织微阵列典型过程的流程图

材料

注意：使用经过 DEPC 处理的 H_2O 配制本实验方案中全部试剂（见第 3 章关于 RNA 操作特别注意事项的信息栏）。

608

组织标本

对照组织

在所有 TMA 中包含已知表型的组织。这些组织可能包括，例如，一栏正常组织

（如肝脏和肾脏表皮），来自明确细胞系的细胞团，或是异种移植的肿瘤标本。已知表型的组织作为内部参照，用于验证后续分子和免疫组化分析的特异性和灵敏度。

组织标本

这里介绍的是使用存档的经过甲醛固定的组织制备阵列。这些组织可能在不同的条件下固定，保存的时间也不相同，这就限制了它们在一些分子分析中的应用。例如，应用未缓冲甲醛固定的组织，这是直到数年之前常用的方法——进行 RNA 原位杂交（ISH）或者是免疫组织化学（ISH），会得到高度衰减的信号。经过长期保存，乙醇固定组织中的 RNA 也容易降解。新鲜或冰冻组织在经过保持核酸的标准方法（如缓冲的甲醛）固定和石蜡包埋后可用于制备组织微阵列。

参考组织切片，裱贴于载玻片上并用苏木精和曙红（H&E）染色。

来自供体组织块（步骤 2）的切片可以染色，用作参考样本。

专用设备

黏性胶带切片系统（Instrumedics Inc.，Hackensack，New Jersey）

双目体视显微镜

包埋用石蜡块（Polysciences）

显微镜用薄片切片机

石蜡（ParaPlast＋）

石蜡包埋工作站

Superfrost Plus 载玻片（Fisher Scientific 或者 Menzel Glaser）

组织阵列排列仪器（如 Beecher Instruments，Silver Springs，Maryland，和 Chemicon International，Temecula，California）

609 附加试剂和设备

本实验方案的步骤 2 需要组织切片染色试剂，包括苏木精（Sigma）和曙红溶液（曙红 Y，1% 水溶液，EM Diagnostic Systems），以及组织样品石蜡包埋的试剂（Spector and Goldman 1998，见第 98 章；Bancroft and Stevens 1996）。

方法

组织标本的准备或选择

1. 选择“供体”组织块，选择那些标本厚度接近一致的组织块。

   用于 TMA 的组织标本块厚度至少为 1 mm，以适合制备组织阵列。组织块的最佳厚度为 3～4 mm，这样可以从阵列蜡块得到更多数量的切片。

2. 为每块供体组织块制备，或者得到用苏木精和曙红（H&E 染色切片）新鲜染色的参考切片（约 3 μm）。

   参考切片用于确认为制备阵列进行取样的合适区域。其他任何最近染色的切片（如免疫染色切片）也可用于这一目的。在回收芯样之前不要从供体蜡块中取得更多的切片，这样做会降低原始参考切片的代表性，并可能带

来取样误差。

3. 检查用供体组织制备的染色切片，并确定哪些组织块中包含有用于取样的代表性区域。
4. 在参考切片上，根据需要尽可能多的标记这些区域。
5. 准备一个空白石蜡块用作组织样品的载体。
6. 将标准包埋用石蜡（熔点 55～58 ℃）熔化并倒入包埋模具中。

   确保模具中没有蜡或者其他残留物。大的包埋模具（30 mm×45 mm×10 mm）比病理实验室常用的 25 mm×35 mm×5 mm 标准模具容易操作。
7. 在熔化的石蜡上方放置一个塑料的蜡块容器。
8. 等到整体冷却后，将模具和固化的载样蜡块分离。（译者注：原文误为 donor block，供体蜡块）

   确保石蜡中没有气泡。
9. 设计 TMA 的布局，使之能容纳下 TMA 中将要排列的供体组织数量。为了方便 TMA 的显微分析，将芯样排布成亚区，每个亚区包含 100～200 个标本。在阵列蜡块的角落处留下足够的空间（2.5～3 mm），以防止石蜡的破碎。

   在尺寸为 30 mm×45 mm 的蜡块中最多可排列大约 1200 个标本。然而，制备包含超过 700 个标本的蜡块需要紧密排列标本，这提高了对排列可靠性的要求。包含 400～500 个标本的蜡块制备与此相比要容易一些。当使用的穿刺针直径为 0.6 mm 时，阵列中两个相邻标本的中心距可以从 0.65 mm 到 1 mm（通常为 0.8 mm）。

610

使用 Beecher 阵列排列仪器制备 TMA 蜡块

制备包含大于 500 个标本的大规模阵列，必须使用精密仪器，以保证所需精度。以下操作依赖于一种在设计上类似于那些 cDNA 阵列制备仪器的装置（图 8-4）。

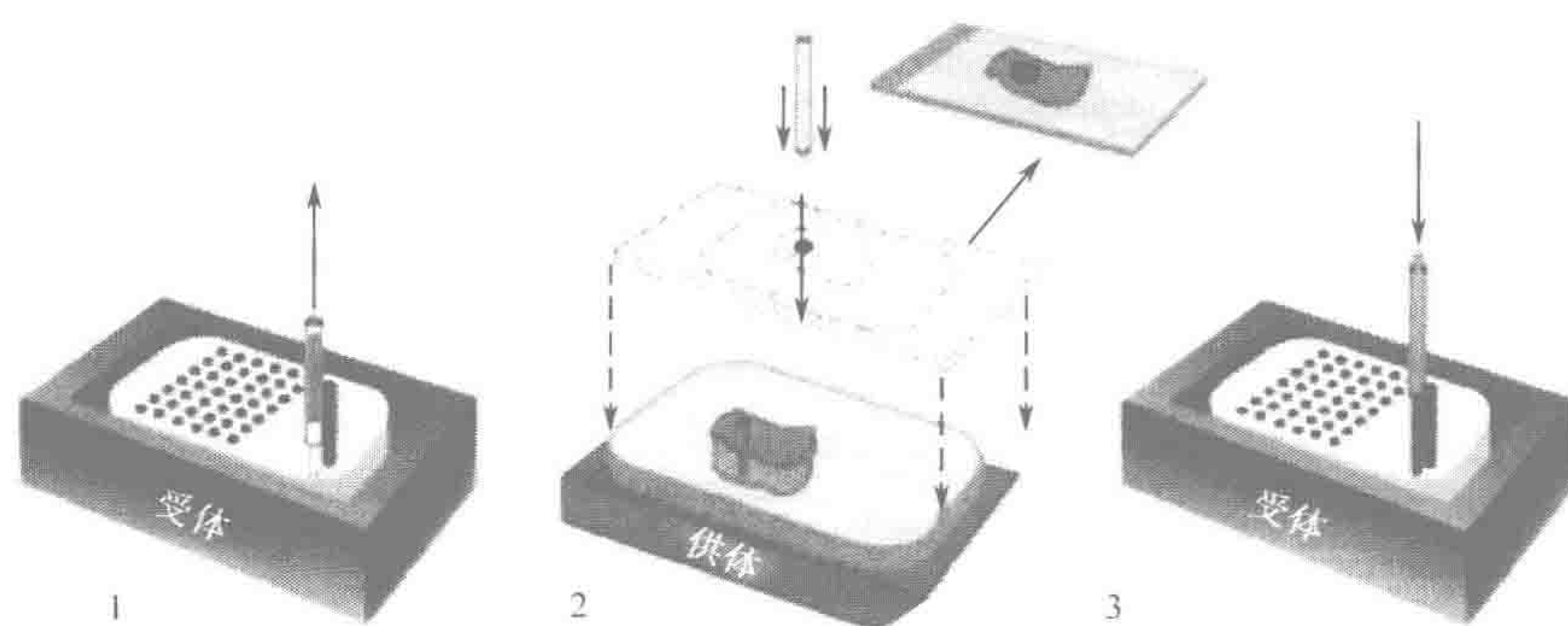


图 8-4 组织微阵列的制备

（1）在阵列上为排列一个新标本钻孔；（2）从供体组织蜡块上具有形态学代表性的位置取样；（3）将组织芯样排列到载体（阵列）蜡块中各自的位置上

10. 鉴别和选择组织上的部位：

    a. 以 H&E 染色的参考切片为向导，确保组织芯样取自先前（步骤 4）已经鉴

别用作分析的区域。有时候，可以将要钻取的区域从“供体”蜡块表面直接辨别和标记。然而，对于钻孔取样的位置常常需要更加精确的加以选取。用双目体视显微镜精确选择或观察 H&E 染色切片上预先标记好的取样位置。

b. 目标区域确认后，移去染色的切片，并按照如下方法取得样品。

11. 使用阵列排列仪器获取组织芯样：

a. 在载样蜡块上钻一个孔 [图 8-4（1）]。

b. 从供体组织蜡块得到组织样品 [图 8-4（2）]。

阵列排列仪器有两种不同尺寸的针，分别用于移除石蜡，以及从供体蜡块获取组织芯样。细针用于在载样蜡块上钻孔。细针的外径略小于粗针的内径。因此，取样（粗针）得到的每个组织样芯可以紧密的安放入载样蜡块的孔中。

c. 将该标本放入载样蜡块上相应的孔中。在载样蜡块 z 轴方向上仔细的安放标本 [图 8-4（3）]。

注意所有样品放置的深度要一致，以保证得到大量切片包含了来自全部标本的代表性组织。如果一个标本在载样蜡块上没用推入足够的深，部分组织会在开始切片之前，削平载样蜡块时损失掉。如果样品推入得太深，它就不会出现在阵列蜡块的最初切片中。

611

阵列排列仪器每次可以夹持住一个载样蜡块，通过位于夹具底部的磁铁精确地锁定在固定位置。在多个分别固定在各自夹具上的载样蜡块之间转换，可以复制多个阵列蜡块。

12. 重复步骤 10 和 11，直到标本向阵列蜡块的转移全部完成。

13. 将阵列蜡块放置于水平表面上，准备切片。为了使样品达到最佳的水平：

a. 将阵列蜡块放入 37 ℃培养箱中 15 min。

b. 加温后，在阵列蜡块上方放置一张载玻片，均匀施压使得阵列中所有组织芯样处在同一平面上。

c. 让阵列蜡块冷却至室温。这时的阵列蜡块就可用于切片了。

14. 对阵列蜡块进行 5 μm 切片，产生若干的 TMA。根据包埋在蜡块中的组织类型，可以将阵列块冷却至－10℃和＋10℃之间。将切片漂浮于灭菌蒸馏水（用于 IHC），或是经过 DEPC 处理的灭菌蒸馏水（用于 ISH 和 IHC）中，然后将切片转移到 Superfrost Plus 载玻片上。

如果仅用少量组织（小于 200）制备阵列，可以使用标准的显微镜薄片切片机对阵列蜡块进行切片。对于大规模阵列，黏性胶带切片辅助使得收集连续的切片，以及将它们精确排列于载玻片表面变得简单易行。由于不同组织切片的附着特性各不相同，整张切片，或者是切片中的单个标本在后续的处理中，特别是在增强通透性（ISH，实验方案 2）或者抗原修复（免疫细胞化学，实验方案 3）过程中会发生分离。胶带切片系统中，涂抹了黏附剂的载玻片极大地改善了 TMA 的附着特性（图 8-5）。

切片厚度为 5 μm 可以改善其黏附特性，特别对于长时间或者严格的冲洗过程，如 ISH，或者荧光原位杂交（FISH）。

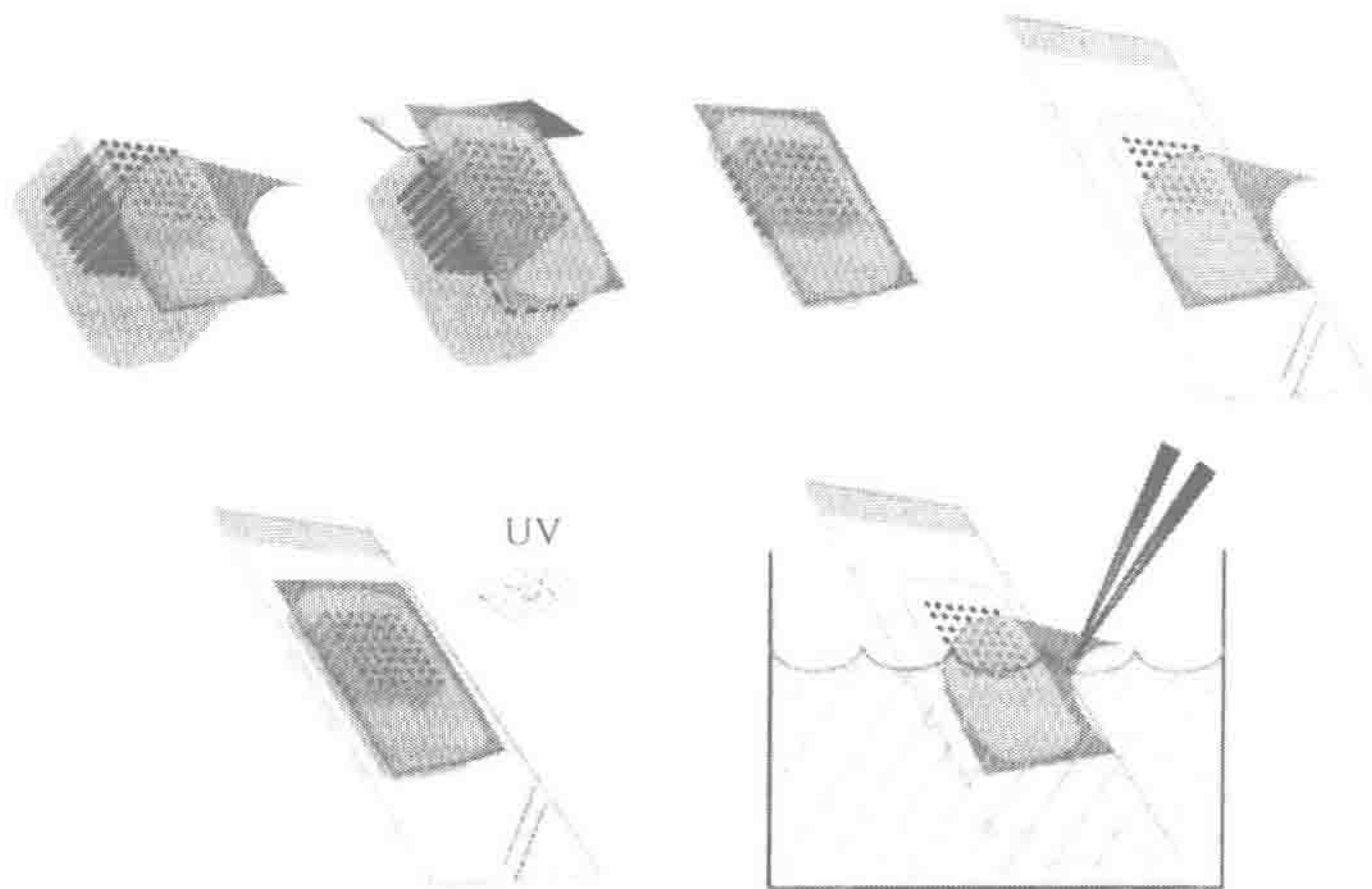


图 8-5 应用黏性胶带切片技术对阵列蜡块进行切片

仔细地将一条胶带放在阵列蜡块的上方。显微镜薄片切片机的刀在胶带下方切片，带有一张 5μm 切片的胶带随后被转移到覆盖了黏附剂的载玻片上。经过短暂的紫外照射使得载玻片上的黏附剂聚合，在溶剂溶液中去除胶带

15. 将 TMA 在 60 ℃烘烤 30 min（用于 IHC）到 2 h（用于 FISH、ISH），除去过量的石蜡和水分，并促进 TMA 和载玻片之间的附着。4 ℃下，TMA 可以在无水环境，如干燥器中保存至少 12 个月。

    用于 ISH 的切片最好戴着手套操作，因为与裸露皮肤的接触可能发生 RNase 的污染，导致 TMA 中目的 RNA 的降解。

质量控制

检查每一个阵列，确定缺失点的百分数，非代表性点的百分数，以及变形或其他损伤点的百分数。关于如何达到高的成功百分比，没有固定的指导方针。然而，为了减少取样偏差的可能性，成功点的百分比要大于 90%，这一目标可通过仔细的阵列制备实现。关于 TMA 其他的期望属性有：高精度的 x 和 y 坐标定位、好的染色特性以及在 ISH（方案 2）和 FISH 中有高的信噪比。

表 8-1 组织微阵列制备的疑难解答指南

问 题	可能原因	解决方案
阵列中一些标本缺失	一些标本在阵列中推入太深	重新排列标本的阵列，注意芯样 z 轴的位置
	一些组织芯样太短	确保用于制备阵列的组织蜡块厚度尽量一致，并且越厚越好
	应用胶带技术切片时，一些标本发生分离	确保胶带和涂抹有黏附剂的载玻片质量很好。在紫外聚合之前，除去任何可能的气泡

续表

问　题	可能原因	解决方案
	在染色/杂交过程中标本分离脱落	使用黏性胶带技术切片
阵列中的标本不具备代表性	取样不当或者随着对阵列蜡块进行切片数量的增加，形态学上发生改变	借助显微镜确认取样来自组织蜡块正确的区域。对多重位置进行取样
阵列上有变形	制备阵列的针发生弯曲	更换所用的针
阵列中间部分向上突起		增加阵列中相邻标本之间的距离和（或）加深阵列蜡块上钻孔的深度。阵列排列完成后，按照前面所述使其表面处于水平
阵列蜡块上的石蜡发生破裂	标本放置得距离阵列蜡块边缘太近	在边缘留 2.5～3 mm 空白。使用较大的载样阵列蜡块
残留的黏附剂干扰分子分析	UV 处理聚合黏附剂的时间太短	对于小规模阵列，无需使用胶带切片方法。使用较长的紫外透射照明时间

613

附加方案：手工制备组织阵列

Karl Brand*（Peter MacCallum 癌症中心，东墨尔本，维多利亚 3002，澳大利亚）

组织阵列的制备是由供体蜡块取得芯样，并重新将它们以紧密间隔的网格样式排列于石蜡中，用于后续的切片和放置在显微镜载玻片上。这一过程可以在特殊设计的仪器（Kononen et al. 1998）（见实验方案 1）的辅助下进行，或者如下所述，通过手工方式将来自供体的芯样转移到载样蜡块中。手工转移可以将最多 200 个标本排列到一张载玻片上，产生对标本数量少，样品尺寸大，或者是具有复杂解剖结构的标本具有很大应用价值的 TMA，如对中枢神经系统或哺乳动物胚胎的切片。

手工制备组织阵列要求的灵活性并不比常规石蜡超薄切片要多。与任何手工技术一样，它有一个学习曲线。通过勤学苦练，使用本实验方案中的详细方法，在第四次制备阵列时，也就是说经过大约 5 h 的练习，就可以得到具有出版质量的阵列。制备阵列的最初尝试可以用 2.0 mm，最好 3.0 mm 直径的芯样，并且每个阵列小于 50 个芯样，然后再过渡到直径 1.0 mm 或更细的较小芯样，它们的处理比较困难。

附加材料

黏性胶带组织切片辅助（Instrumedics，Hackensack，New Jersey）（可选）
对照组织、实验材料、供体蜡块和参考样本（见主材料清单）
生物芯样活检仪器（Stiefel Laboratories Inc.，Coral Gables，Florida 或者 KaiIndustries Co.，Ltd. Oyana，Cifu，Japan）
切片漂浮水浴
铁针（定制的或者是回形针已足够）
96“U”型孔板（Greiner）

* 当前地址：Erasmus University，3015 CE，Rotterdam，The Netherlands.

方法

准备

1. 设计组织阵列的样式和内容。
2. 得到关于供体蜡块的染色切片，或者切取 3μm 参考切片并用 H&E 染色切片。
3. 用永久记号笔勾勒出想要取样的部位。

获取芯样

4. 将供体蜡块已经切片的表面朝上，放置在平的，没有振动的表面，如工作台上。
5. 将标本蜡块与 H&E 参考切片对准，并将芯样活检设备放在选定的取样区域上方。
6. 在不影响芯样钻取仪器与供体蜡块相对位置的前提下，移去 H&E 参考切片。
7. 不加耽搁，将芯样钻取仪器插入供体蜡块，直至它从支持盒的底部“穿出”。

    沿垂直方向加力以确保芯样设备笔直向下穿出供体蜡块。过度用力可能导致对供体蜡块和（或）芯样不必要的损伤。当取样区域密度有变化，或者取样位置紧靠先前移取芯样形成的孔穴时，力的方向（相对于垂直平面）可能需要改变。任何方向的改变都要避开先前钻取芯样的部位。

8. 轻轻地从供体蜡块中退出芯样活检设备，通过扭转运动克服阻力。
9. 用退出针将芯样推入到 96“U”型孔板上预先指定的位置。

    96“U”型孔板是非常好的芯样候用保存器具，因为它们有合适的网格形式和尺寸（7 列×8 行）。从 96“U”型孔板取回芯样要比从平底板容易些。

    推出芯样时要仔细。退出针会加压于后续用于切片的芯样表面。过度用力会损伤包埋在蜡块中的组织。

    每种组织阵列使用一个 96“U”型孔板。重复的阵列可以同步制备，将来自同一供体蜡块的芯样分别放入各自的 96“U”型孔板。

614

芯样阵列排列

10. 准备一个有足够长度（至少 25 mm）、宽度（至少 30 mm）、和深度（至少 4 mm）的包埋盒，用来装芯样。
    a. 用双面胶带覆盖模具的底部（后来形成组织阵列的切片表面）。确保胶带与模具的 x 轴或者 y 轴对准。胶带作为放置芯样的向导。
    b. 用永久性记号笔标出包埋模具的中心位置。这有助于阵列排列时正确的对准芯样位置（见步骤 11 的注解）。

    避免过度触摸双面胶带的任何一面。如果有可能，使用一侧覆盖了保护贴纸的双面胶带。将胶带未覆盖贴纸的一面粘贴到模具底部，并在排列芯样阵列之前除去覆盖的保护纸。

11. 用锐利的镊子从 96“U”型孔板夹取芯样，将芯样已切片的端面浸入熔化的石蜡，并迅速将芯样放置到包埋模具中合适位置的双面胶带上。对每个芯样重复该操作。最终目标是在包埋模具中生成若干行和列的芯样，并且与“U”型孔板上的行和列对应。

不要将第一个芯样放置在“A1”位置，该位置空缺留作定位用。而是将芯样放在阵列最顶上一行的中间位置（例如，一个具有7列的阵列的A4位置）。包埋模具的中心位置在先前的准备步骤中已经预先标记过。仅仅对于第一行（行A），从该“顶部-中心”位置逐步回填。这时，行A用作排列剩余芯样的精确向导模板。

将芯样浸入熔化的石蜡有助于芯样与双面胶带的黏附。

如果石蜡冷却太快，或者芯样位置排列不正确，仔细去除黏在芯样表面冷却变硬的石蜡并重复这一操作。

确保芯样的长度不要超过包埋模具的深度。尺寸过大的芯样在包埋过程中可能会碰倒。如果有必要，将盒子底部盖到包埋模具上，检查芯样的长度。仔细除去任何与盒子底部接触的芯样，并修平多余的石蜡。将已切片端面的对侧用于修平。重新放回芯样并检查其长度，确保它与盒子之间的空隙。

注意：一些种类的石蜡含有二甲基亚砜（DMSO）。操作熔化的石蜡时要戴上手套。

包埋

12. 将包含放置好芯样的模具加热20 min到42 ℃。

    加热减慢了熔化的石蜡与芯样接触时降温和固化的过程（步骤13），从而减少了气泡的出现。

    重要注意事项：此时要更加仔细操作以避免干扰排列好的芯样。这种干扰对于后续鉴定芯样可能是损失惨重的。

13. 轻轻地将熔化的石蜡倒入排列了芯样的包埋模具。必须非常仔细以避免干扰芯样。继续倒入石蜡直至充满模具为止。

    控制石蜡流动的方向在角落处并且距离排列的阵列尽可能的远。将熔化的石蜡直接倒到芯样上增加了干扰排列好阵列的可能性。

14. 将一个盒子放在包埋模具上面（和常规包埋一样），并迅速在包埋模/盒子上方加入一层附加的石蜡。

    模具中石蜡加入不足就会导致石蜡块中孔穴的形成，从而在切片时带来问题。

15. 将包埋模具/盒子在冰上放置30 min使石蜡变硬。

    不要将模具放置在0 ℃以下。暴露在低温会使得石蜡变硬时破裂。

16. 用螺丝起子或者其他钝刀器具，借助杠杆原理轻轻的使得包埋模具和盒子分离。修平多余的石蜡。

17. 按照主要实验方案中步骤14和步骤15所述，进行标本切片和烘片。

疑难解答信息栏

• **气泡** 将芯样包埋到熔化的石蜡中时，蜡中经常会形成气泡。一到两个气泡通常是可以接受的，并且不会给切片带来困难。如果产生的气泡过多，可以在切片时向空腔中填入石蜡来消除影响。具体做法是将组织阵列蜡块加热到50 ℃，然后向空穴中直接注入熔化的石蜡。手工将石蜡轻压到空穴中。切片前将阵列蜡块置于室温下冷却。

在芯样包埋到熔化的石蜡之前，将组织阵列预热到42 ℃，可以减少气泡的形成。将预热温度提高到46 ℃可以进一步减少气泡的发生率。预热温度不能大于46 ℃。在较高的温度下，把芯样固

定到双面胶带上的石蜡的黏附特性下降，在用熔化的石蜡包埋时，会导致芯样从与包埋模的系泊处脱离。这里推荐的温度适合于熔点55～56 ℃的石蜡（如ParaPlast+）。

• **供体蜡块中组织的破碎**　破碎可能是由于组织固定得不好或供体蜡块上的芯样孔数量过多。确保那些打算用于阵列制备的组织固定的时间足够长，以避免芯样的破碎。修复破碎的供体蜡块和填充芯样孔可以按照如下的方法操作：

1. 预热标本供体蜡块到50 ℃。这有助于把软化的供体蜡块和熔化的石蜡"混合"。
2. 将熔化的石蜡直接注入芯样孔。手工将石蜡轻压到组织碎片/孔穴中以加大穿透深度，并重新包埋组织标本。
3. 切片前让蜡块冷却到室温。

• **黏附/组织缺失**　由于芯样切片尺寸很小，因而很容易从组织阵列切片分离。为了减少这一问题，可以使用Supcrfrost Plus载玻片，并确保脱蜡前组织至少烘烤2 h。采取以下方法可以进一步防止芯样切片的缺失：

1. 在孵育过程中，使用高出组织表面的盖玻片（如Lifter Slips，Erie Scientific Company Portsmouth，New Hampshire）。
2. 用塑料片替代盖玻片。保存尼龙膜的塑料（用于Southern和northern印迹法）可用作理想的材料。所使用的塑料片必须完全没有灰尘和污染物。
3. 分两个阶段烘烤载玻片，每次在60 ℃下连续2 h。在第一次烘烤后，让载玻片在室温下冷却1 h后再进行第二次烘烤。要注意延长烘烤时间会导致形态学上的破坏，并可能对RNA的完整性不利。只有在其他各种组织黏附方法都没有效果的情况下，才能加长烘烤时间。
4. 使用其他组织黏合剂（如硅烷化和涂有明胶的载玻片、黏性胶带和来自Instrumedics，Flackensack，New Jersey的组织切片辅助物）。

方案2

616

应用组织阵列进行原位杂交

Karl Brand*（Peter MacCallum癌症中心，东墨尔本，维多利亚3002，澳大利亚）

原位杂交技术（ISH）用于在组织切片上观察核酸的分布（John et al. 1969；Pardue and Gall 1969）。本实验方案介绍一种非放射性的，利用地高辛（DIG）标记探针检测组织微阵列中mRNA的方法。这项技术是建立在20世纪90年代早期发展起来的一种方法（Heiles et al. 1988；Holtke and Kessler 1990；Kessler et al. 1990；Muhlegger et al. 1990）的基础上，首先是Boehringer Mannheim，近来由Roche Inc. 将其商业化。关于使用DIG进行核酸标记和检测DIG配体方法的背景知识，请参看Sambrook and Russell（2001，A9.38-9.42页）。

用于ISH组织的固定

固定方法的类型、温度和固定时间的长短，对用于ISH组织样品的后续处理都有

* 当前地址：Erasmus University，3015 GE，Rotterdam，The Netherlands.

很重要的影响。固定时蛋白质交联/凝结的程度越大，靶核酸就越容易受到掩蔽。这一问题对与组蛋白和其他蛋白质紧密结合的基因组 DNA 来讲，要比 RNA 更为严重，后者与蛋白质的结合更加短暂，组织化程度也较低。用于 ISH 来分析 mRNA 表达的组织，推荐的固定方法为使用缓冲的甲醛（pH 7.5），在室温下固定 16～26 h，这取决于样品的尺寸（Bancroft and Stevens 1996）。

这里介绍的 ISH 实验方案是针对经过甲醛固定和石蜡包埋的组织而进行优化的。该方案经过改进后可以应用于其他方法固定的组织，但是采用了不同方法固定的组织集中于同一个阵列中时，可能会带来问题。在此情形下，对来自相同组织阵列的一系列切片中的每张切片，都可能需要一种不同的实验方案（请见蛋白酶 K 处理暴露靶 mRNA，下面）。很显然，这造成了来自同一阵列蜡块的不同切片中，对靶 RNA 的检测效率可能出现系统性的偏差。为了防止这些问题发生，一个阵列蜡块中的全部芯样最好要采取同样的固定方法进行固定。

617

蛋白酶 K 处理暴露靶 mRNA

组织制备的一个关键步骤是，除去与 mRNA 结合的，由蛋白质经化学固定后形成的交联。在整个 ISH 操作过程中，固定产生的交联在数量上要尽可能的少，以使得标记探针易于接近，同时要保留足够的细胞结构，让靶 RNA 保存在其正常的位置。实现这些目标的最佳方法是用蛋白酶，如蛋白酶 K 处理组织切片，这类酶有广泛的底物特异性，并且有可以失活组织切片中核酸酶的附加优点。

ISH 中去掩蔽过程的优化，主要是通过调节反应混合物中蛋白酶 K 的浓度，改变蛋白水解消化的程度。对于每种探针和组织阵列，其最佳的消化条件要根据经验来确定。那些用甲醛固定的组织，建议从 1 μg/ml 蛋白酶 K 开始（在 37 ℃孵育 30 min）。然而，对于致密的组织，如子宫组织，所需要的酶浓度可能高达 15 μg/ml。如上所述，由不同方法固定的组织或超过一种类型组织的阵列，可能需要在各种不同条件下进行消化。例如，包含了乳房、肌肉和来源于软骨的组织的阵列，可能需要一系列蛋白酶 K 浓度不断升高（1～5 μg/ml）的独立消化处理，以在全部三种组织中获得对 RNA 的最佳去掩蔽效果。

探针设计

一个设计优良的反义 RNA 探针，是与感兴趣的靶 mRNA 进行特异而高效杂交的关键。两个相关联基因家族间保守性最低的区域，通常是在 mRNA 的 3′-非转录区发现（Lipshutz et al. 1999），并且探针主要也是指向靶 mRNA 的这一片段。对寡核苷酸探针而言，DIG-标记的转录片段长度为 25～400 bp，长度为 25 个核苷酸的寡核苷酸探针在包含了 8 种或更多种探针的混合物中可以很好的工作（Komminoth 1992；Komminoth et al. 1992）。随着探针长度的增加到 400 个核苷酸以上时，其穿越固定的组织与靶 mRNA 结合的能力受到的限制不断增加，信号程度也随之成比例的衰退。较大的 RNA 片段必须通过有限的碱水解以减小尺寸（Cox et al. 1984；Angerer et al. 1987；Bales et al. 1993）。

ISH 中与 mRNA 结合的反义 RNA 探针是在噬菌体 T7 依赖于 DNA 的 RNA 聚合

酶催化下，通过体外反应合成的，使用DIG-标记的核苷酸部分取代标准的rNTP（Holtke and Kessler 1990；Komminoth 1992；Komminoth et al. 1992）。关于该反应，以及其他合成与纯化DIG-标记反义探针的方法的讨论，请见Sambrook and Russell（2001）。反应中使用低浓度的修饰核苷酸，以确保每个探针分子中仅掺入少量的标记分子（每20～25个聚合的核苷酸中约有1个标记的核苷酸）。修饰核苷酸掺入过多会降低探针与靶mRNA的杂交效率。只有当少数碱基被其标记的对应物取代时，杂交后抗体对DIG-取代的RNA识别才最为有效。

对照

618

阴性对照

阴性对照探针（negative control probes，NCP）对确定观察到的信号是由于实验探针与其靶序列的特异结合，还是由非特异性吸附引起的极为重要。NCP与反义的实验探针在长度和GC含量上应该保持一致。常用的阴性探针有两类：正义阴性对照和门间阴性对照。

- **正义阴性对照探针**（sense negative control probe，SNCP）是来自正义链，也就是说，序列上与反义探针完全互补的RNA链。探针序列主要是亚克隆到在插入区域两侧带有T3和T7转录位点的载体中。理想情况下，SNCP与细胞中其他任何转录产物不发生结合。然而在实际应用时，全部SNCP中约10%显示出中等程度的染色，并且这种染色常常表现为高特异性。为将这一问题降至最低，反义和正义探针的设计要特别仔细，避免与组织阵列中其他可能存在的RNA同源。最好在合适的数据库中认真仔细搜寻与反义探针和其正义对应物互补的那些序列是否存在。
- **门间阴性对照探针**（interphylum negative control probes，INCP）与来自无关的门的特定靶序列直接对应，为关于SNCP的警示性说明提供了解决方法。直接对应于病毒或植物序列的探针，与哺乳动物序列有效杂交的几率很低。在设计INCP时，要通过数据库的检索来识别和减小INCP和组织阵列中可能出现的mRNA转录物之间任何潜在的交叉杂交序列。

阳性对照

阳性对照探针（positive control probes，PCP）用于检测ISH实验过程中可能出现的技术问题，也用于估计ISH实验之间的变异程度。当与阳性对照组织共同使用时，PCP也可以作为对ISH实验方法效率的一种测量。任何普遍性表达的管家基因，以适当水平表达时，就可以用作阳性对照，如编码细胞角蛋白的基因。

标记对照

标记对照探针（labeling control probes，LCP）用于检查转录反应之间标记效率的变化。使用PCR序列作为LCP，可以测量不同ISH实验之间的变化。PCP和LCP分别由不同的转录反应标记：PCP是在ISH技术建立的时候，而LCP则是在任何新的实验探针被标记时。标记效率的差别可以通过PCP和LCP对阳性对照组织染色的强度和分布的比较来测量（见下文）。

组织对照

常用的组织对照有两种类型：阳性对照组织和用于验证 NCP 的阳性对照组织。

619

- **阳性对照组织**（positive control tissue，PCT）该组织被认为，或者先前已知道其中 mRNA 的表达水平，可以很容易被用作 PCP 的探针检测到。一个 PCT 可以提供每种 PCP 的染色图谱。
- **用于验证 NCP 的阳性对照组织**（positive control tissue for verification of NCP）该组织用于验证 NCP 的特异性和功能。必须要有一种这样的对照，来确证靶组织未被染色不是由标记较差，或者探针检测效率低的原因造成的。特别是对于正义 NCP，并不总是可以获得被认为，或先前已知道其中会表达与 NCP 互补的 mRNA 的组织，一种可能的替代方法是应用经适当质粒转染后会表达 NCP 序列的细胞系。

杂交条件

ISH 的杂交条件是建立在理论和实践经验两者综合的基础之上。三个主要的变化因素是杂交温度、溶剂和探针的浓度。

- **杂交温度**　每种探针在一系列组织阵列上的温度要根据经验来确定。每当第一次使用，或者“优化”一种探针时，杂交要在 42～60 ℃进行。
- **溶剂**　在这里提供的实验方案中，用于阵列的杂交混合物包含了 60%的甲酰胺，孵育盒也用含有 40%甲酰胺的溶液湿润。在 16 h 孵育过程中，杂交缓冲液中甲酰胺的浓度会缓慢下降。根据经验，这种 60/40 缓冲液/孵育盒的甲酰胺梯度，用于大多数探针时可以得到理想的结果。这种情况在原理上与“热启动”和“下降”的聚合酶链反应（PCR）相似，探针与靶序列的复性发生在高度严格条件下，并在杂交中逐渐减弱。然而有一些探针，得益于应用了 40/60 缓冲液/孵育盒梯度，使得杂交过程中杂交缓冲液的甲酰胺含量上升。
- **探针浓度**　探针与靶序列的复性是一个拟二级反应。因此，初始探针浓度越高，探针与靶的复性速度越快（Wetmur and Davidson 1968；Wetmur 1991）。由于饱和大量靶 mRNA 需要的探针浓度约为 300 ng/ml 杂交缓冲液，每千个探针碱基序列（Wilkinson 1992），使用探针浓度大于 300 ng/ml/kb 不会增加靶信号，但会使得非特异性背景增高。当靶 mRNA 的拷贝数量很少时，可以使用低浓度的探针。在所有情形下，每种探针的最佳浓度要根据经验来确定。

杂交后处理

620

RNase A

杂交后用 RNase A 处理组织除去未杂交的单链 RNA。探针与其靶 mRNA 之间形成的杂交体可以抵抗消化并保持完整。如果一切正常，RNase 处理应该可以消除非特异性背景而不会降低特异性信号的强度。事实上，特异性和非特异性信号常常都会显著减弱（Yang et al. 1999）。信噪比的净增益通常不大。因此，只有当非特异信号居高不下，而信号又是足够的强并且可以补偿信号损失时才使用 RNase A。

信号放大

应用免疫学方法检测DIG标记探针时，能够对杂交信号进行放大（Bobrow et al. 1992）。一种常用的放大方法是，使用在辣根过氧化物酶（horseradish peroxidase，HRP）催化下的、从tyramide-biotinyl（译者注：参看TSATM，Tyramide Signal Amplification，http://www.nenlifesci.com）到二硝基苯酚（dinitrophenyl，DNP）的转化。DNP沉淀物沉积在标记探针所在的位置，通过碱性磷酸酶缀合的抗DNP抗体（anti-DNP-AP）来检测。应用最佳的ISH实验方法时，几乎不需要对信号进行放大。然而，当靶mRNA数量稀少，固定不好，或者靶mRNA的暴露不完全时，信号放大可以增强信号和改善信噪比。我们还是推荐只有当非放大式ISH方法已经完全得到优化，而且经过信号放大后信噪比的增加得到证实时，信号放大才作常规使用。除标准的ISH方案之外，信号放大还包含如下步骤：

- 组织预处理时淬灭内源的过氧化物酶活性。
- 使用为通常浓度1/4的探针。
- 通过HRP催化的DNP沉淀放大信号。
- 信号放大后进行充分洗涤。

本章叙述的放大步骤是根据New England Nuclear发展的技术［TSA Plus DNP（AP）体系］。

ISH优化的一般策略

表8-2列举了建立ISH实验时要加以优化的，和应用任何新的探针或者组织类型时必须重新优化的参数。请注意：为每步操作给出的实验条件的建议范围，是为经过甲醛固定和石蜡包埋的组织而设计的，可能不适合于使用其他方法固定的组织。

表8-2　DIG-ISH的优化

参　数	改变的原因	建议的优化范围[a]
预杂交		
蛋白酶K消化暴露mRNA	不同的组织；不同的固定方法；探针染色不充分	1 μg/ml、5 μg/ml，和10 μg/ml（1 μg/ml）
探针浓度	不同种类/新探针	100 ng/ml和200 ng/ml（200 ng/ml） 25～100 ng/ml（放大的ISH）（50 ng/ml，放大的ISH）
杂交温度	不同种类/新探针；不同的组织；不同的固定方法	42 ℃，50 ℃和60 ℃
杂交后		
RNase处理	仅当可以消除过高的背景时	20 μg/ml RNase处理（未进行操作）
抗DIG-AP抗体孵育	探针信号不足；背景染色过高	1∶500～1∶1500（1∶1000）
抗DIG-HRP（放大式ISH）	优化放大参数	1∶50～1∶500（1∶100）
抗DNP-AP（放大式ISH）	优化放大参数	1∶25～1∶500（1∶50）
DNP放大底物	优化放大参数	5～30 min（10 min）

a 括号中给出了建议的初始值。

RNA 的操作

DIG 标记的 RNA 会被很快降解，而且事实上在实验室中到处存在的 RNase 很难失活 。完全消除 RNase 是不可能的，也是不必要的。然而，在制备用于杂交的组织时，以及杂交过程中，还是推荐将 RNase 活性降至最低，以防止杂交信号的减弱。关于在实验过程中减弱 RNase 活性方法的详细讨论，请参看本册书第 3 章关于 **RNA 操作的注意事项**的信息栏，以及 Sambrook 和 Russell (2001，第 7 章)。简要地说，应采纳下列的注意事项，将 RNase 活性降至 DIG-ISH 可接受的水平：

- 操作任何将要与探针直接接触的材料，如载玻片、载玻片架、盖玻片、微型离心管以及预杂交和杂交的容器时，要套上乳胶手套。
- 使用经过 DEPC 处理的 H_2O 和试剂。
- 所有玻璃器具使用前，在 180 ℃ 烘烤 60 min。
- 预留一批专用的实验室器具，专门用于预杂交和杂交溶液的制备与操作。

材料

注意：标有〈!〉材料的正确操作方法见附录 3。

缓冲液和溶液

关于储存液、缓冲液和试剂的成分见附录 2。

将储存液稀释至合适的浓度。

乙酸酐 (0.25%)〈!〉

封闭溶液 (Roche)

50×Denhardt 溶液

10×DIG RNA 标记混合物 (1 277 073，Roche)

二硫苏糖醇 (dithiothreitol，DTT) (250 mmol/L)〈!〉EDTA (250 mmol/L，pH 8.0)

乙醇 (无水化学和分析纯)

溴化乙啶〈!〉

甲酰胺 (去离子的)〈!〉

HCl (0.2 mol/L)〈!〉

Histolene (Pathtech Diagnostic)

氯化锂 (Lithium chloride，LiCl) (4mol/L)

顺丁烯二酸缓冲液 (pH 7.5)

100 mmol/L 顺丁烯二酸〈!〉

150 mm NaCl

$MgCl_2 \cdot 6H_2O$

封固剂

见步骤 38

Clearmount（Zymed），Histomount（Zymed），and Kaisers 甘油凝胶（Merck）都是方便的水溶性封固剂。

NBT/BCIP 缓冲液（四唑氮蓝/5-溴-4-氯-3-吲哚基磷酸盐）

100 mmol/L Tris（pH 9.5）

50 mmol/L $MgCl_2$

100 mmol/L NaCl

NBT/BCIP 显色底物（1681451，Roche）〈!〉

10 μl NBT/BCIP 缓冲液

200 μl of NBT/BCIP 溶液

10×磷酸盐缓冲液（PBS）

预杂交/杂交缓冲液

甲酰胺（去离子的）〈!〉	23.8 ml
1 mol/L Tris（pH 7.4）	0.95 ml
250 mm EDTA（pH 8.0）	0.19 ml
5 mol/L NaCl	3.00 ml
葡聚糖硫酸酯	4.80 g
50×Denhardt's 溶液	0.95 ml
H_2O（经 DEPC 处理，灭菌，蒸馏过的）	6.35 ml
总体积	40.0 ml

该缓冲液最好在使用前配制。此外，该缓冲液也可以配置后，分装成小分量在－20 ℃保存最多 3 个月，仍可以得到满意的结果。该缓冲液解冻后不能再次冰冻。

核糖核酸溶液

鲱精 DNA（10 mg/ml）	1 ml
tRNA（25 mg/ml）	1 ml
酵母总 RNA（25 mg/ml）	1.25 ml
H_2O（经 DEPC 处理，灭菌，蒸馏过的）	748 μl
总体积	4 ml

623

将溶液分装成 500 μl 的等份。

RNase 抑制剂（Roche 试剂盒 1 175 025 或单个 799 025）

氯化钠（5 mol/L）

三乙醇胺（0.1 mol/L）〈!〉

Tris（1 mol/L，pH 7.4，pH 8.0，and pH 9.5）〈!〉硫代硫酸钠（10%）20×SSC

酶和酶缓冲液

蛋白酶 K（10 mg/ml）（745 723，Roche）

10×蛋白酶 K 终止液

2% 甘氨酸 PBS 溶液

1×蛋白酶 K 消化缓冲液

100 mm Tris（pH 8.0）

50 mm EDTA (pH 8.0)

定容到1 L，并高压灭菌。

RNase A (109 169，Roche)

T3 RNA 聚合酶 (20 单位/μl) (1031 171，Roche)

T7 RNA 聚合酶 (20 单位/μl) (Roche 试剂盒 1 175 025 或单个 881 775)

10×转录缓冲液 (Roche 试剂盒 1 175 025 或与噬菌体编码的 RNA 聚合酶一起提供)

抗体

抗 DIG-AP Fab′ 片段 (1 093 274，Roche)

抗 DIG-HRP (1 207 733，Roche)

核酸和寡核苷酸

对照用 DIG 标记的 RNA (1 585 746，Roche)

鲱精 DNA (10 mg/ml) (1 467 140，Roche)

模板 DNA

见步骤 1。

tRNA (25 mg/ml) (109 525，Roche)

酵母总 RNA (25 mg/ml) (109 223，Roche)

专用设备

盖玻片 (24×50 mm/60 mm，Biolab Scientific)

载玻片 (Superfrost Plus，Lomb Scientific)

Hybond N^+ 过滤膜 (Amersham)

杂交恒温箱

恒温箱的容积要足够大，以便容纳杂交装置。温度在 40～60 ℃，精度为±1 ℃。使用三个恒温箱，可以同时在三个不同的杂交温度进行 ISH。

杂交支持物

这些支持物使得玻片可以有序的排列在气密杂交盒中的用于保持湿度的甲酰胺缓冲液储藏槽上方。如果没有气密的杂交盒，也可以用塑料包装纸密封杂交支持物。

切片机和组织切片漂浮水浴

必须具备切片机，用于精确的对组织和阵列进行 5 μm 厚的切片。水浴的温度在 20～50 ℃，误差为±1 ℃。

PAP 笔 (DAKO S2002) 用于给组织阵列画圈。

Secureline Linemarker Pen Ⅱ (可以耐受溶剂；Lomb Scientific 1451 Precision Dynamics Corporation San Fernando，California) 用于在载玻片上标记细节。

振荡水浴

水浴尺寸要足够大，可以容纳杂交后对载玻片进行染色的染槽，温度在 37～60 ℃，误差为±1 ℃。

载玻片染色槽和可转移的染色架

存放为杂交做准备和杂交过程中所用载玻片的全部容器，在使用前必须用3% HCl/70% 乙醇彻底洗涤，然后再用经过灭菌、蒸馏和 DEPC 处理的水漂洗两次，进行清洁。根据要处理载玻片的数量选择合适的容器。使用这里介绍的 ISH 步骤，最多可以对 60 片载玻片作常规处理。

载玻片的脱蜡和再水化需要 5 套容器，组织的预处理需要 5 套，而杂交后处理则需要 7 套。当处理的载玻片数量小于 15 片时，可以使用 Coplin 广口瓶。

TSA-Plus DNP (AP) 系统 (NEL746A，New England Nuclear)

U 型微孔板

附加试剂

本方案的步骤 15 需要方案 1 中列举的试剂。

方法

用 DIG 标记 DNA

1. 设计、制备和纯化 DNA 模板，用于在噬菌体 T7-DNA 依赖的 RNA 聚合酶催化下，进行体外转录合成 RNA 探针。

   微量的乙醇或其他杂质会降低标记效率，模板数量不足也会如此。因此，要确保 DNA 模板经过准确的纯化和精确的定量。

2. 在灭菌过的 1.5 ml 微量离心管中将下述试剂和模板 DNA 混合，制备主混合物。将显示的体积乘以（$N+1$），其中 N=需要的转录反应次数。要包括一个包含了阳性对照 DNA 模板的反应，以生成标记对照探针（LCP；见本实验方案的介绍）。

10×转录缓冲液	2 μl
10×DIG 标记混合物	2 μl
RNase 抑制剂 (20 单位/μl)	1 μl
RNA 聚合酶	2 μl
线性 DNA 模板	1 μg
H_2O 灭菌，蒸馏，无 RNase	定容到 18 μl
T7 RNA 聚合酶 (20 单位/μl)	2 μl
总体积	20 μl

   轻敲试管外壁，使试剂混合。将试管离心 1～2 min，使全部液体集中于底部。反应在 37 ℃水浴中孵育 2 h。

   如果标记效率存在问题，要考虑重新纯化 DNA 模板，按比例增加反应规模，和（或）延长反应孵育时间（最多 16 h）。当有关模板的序列信息有限时，后两个量是仅有的选择，这常常是从表达序列标签（expressed sequence tag，EST）生成探针。

   当使用的线性 DNA 是从带有不同启动子的质粒制备时，要加以注意。例如，pBluescript，在多克隆位点任一侧带有噬菌体 T3 和 T7 启动子。如果使用错误的 RNA 聚合酶，会浪费大量的时间和精力。

3. 向反应混合物中加入下列试剂终止反应和沉淀 DIG 标记的转录产物：

250 mmol/L EDTA (pH 8.0)	2.0 μl
4 mol/L LiCl	2.5 μl
绝对乙醇（预冷至－20 ℃）	75 μl

RNA 会从含有 LiCl 的溶液中有效析出。DNA 模板留在溶液里，并在步骤 4 中丢弃。

这里的探针可以在－70 ℃下最多保存 12 个月。

4. 将乙醇混合物在－70 ℃下放置 30 min，然后在小型离心机中，4 ℃下以最大速度离心 15 min 收集沉淀的核酸。除去上层清液，用 75 μl 预冷至－20 ℃的 70% 乙醇清洗 RNA 小团。

要仔细操作，防止小团变松。RNA 小团很容易脱落，随同上清液一起被除去。

从这里开始，对 RNA 探针的操作始终要在工作台上有冰块的条件下进行，并在－70 ℃下保存。

5. 在真空中对小团进行简单的干燥，直至最后一丝可见的上层清液刚刚消失。轻轻振荡，将 RNA 小团重新悬浮在 100 μl 经 DEPC 处理的 H_2O 中。向重新悬浮的探针中加入 150 单位胎盘 RNase 抑制剂，以延长保存时间。

本试验方案中介绍的 ISH 方法中使用了含有葡聚糖硫酸酯的杂交缓冲液，它在溶液中强烈水化，因此减少了混合物中其他分子可利用的溶剂。这会导致表观探针浓度的上升。另一方面，微量乙醇与探针共存会使得探针析出，这会显著增加背景染色。因此要确保全部乙醇都从标记的探针中去除。当所有乙醇被除去以后，cRNA 小团用裸眼几乎是看不见的。

测量 DIG 标记的转录产物浓度

6. 将 DIG 标记探针与已知浓度的标准物（可以购买得到）进行比较，估计它的浓度。

   a. 吸取 18 μl 蒸馏水和 2 μl 制备的探针，加入到 0.5 ml 的 PCR 管中。将管中物质充分混合。

表 8-3　连续稀释摘要

标号	稀　释	加入体积	对照 RNA (100 ng/μl) [最终浓度 (ng/μl)]
1	1/10	2 μl (探针) + 18 μl H_2O	10
2	1/20	10 μl 来自试管 1 + 10 μl H_2O	5
3	1/40	10 μl 来自试管 2+10 μl H_2O	2.5
4	1/80	10 μl 来自试管 3+10 μl H_2O	1.25
5	1/160	10 μl 来自试管 4+10 μl H_2O	6.25
6	1/320	10 μl 来自试管 5+10 μl H_2O	3.11
7	1/640	10 μl 来自试管 6+10 μl H_2O	1.56
8	1/1280	10 μl 来自试管 7+10 μl H_2O	0.78
9	1/2560	10 μl 来自试管 8+10 μl H_2O	0.39
10	1/5120	10 μl 来自试管 9 + 10 μl H_2O	0.19
11	1/10 240	10 μl 来自试管 10 + 10 μl H_2O	0.10

b. 用移液器向 10 个 0.5 ml PCR 管中各加入 10 μl 蒸馏水。这样就得到 1∶20、1∶40等等比例的稀释（表 8-3）。

c. 将 10 μl 稀释的探针加入第一支含有 10 μl H_2O 的试管中，并充分混合。重复这一操作，得到探针和用做标准的商品化 DIG 标记 RNA（100 ng/μl）的 11 个连续双倍稀释。请参见表 8-3 中的连续稀释摘要。

这一操作也可以用 96 孔板很方便地进行，特别是需要对多个样品进行定量时。

7. 以网格方式将每种 1 μl 稀释液点样到尼龙膜上，并使得稀释后每个样品溶液可以与对应稀释的商品化标准品很容易进行对比。在微波炉（800 W）中照射 2 min，或者用 254 nm 紫外光（1.4 J/cm^2）照射 100 s，使得 RNA 与膜结合。
8. 将膜在顺丁烯二酸缓冲液（pH 7.5）中漂洗 30 s。
9. 室温下，将膜在最少量的封闭溶液中搅动 10 min，以封闭未结合的位点。然后加入交联了碱性磷酸酶的抗 DIG 抗体，使得最终的稀释比例为 1∶1000。室温下将溶液孵育 30 min 并轻微搅动。
10. 室温下，将膜在顺丁烯二酸缓冲液（pH 7.5）中漂洗 5 min 并轻微搅动。再重复该操作 2 次，然后将膜沥干。
11. 加入足量的 NBT/BCIP 显色底物，以覆盖整个膜，然后在黑暗中显色，不用振荡。持续显色 30～45 min，直到标准物和样品 DIG 标记 RNA 的稀释物用裸眼都清晰可见。将膜在水中漂洗 2 次，每次 5 min，然后在空气中干燥。

   步骤 8～11 可以很方便地在塑料称量皿中进行。使用可使膜展平的最小尺寸称量皿。确保加入足量的封闭剂/抗体/显色底物溶液，以完全覆盖整个膜。

   注意：NBT/BCIP 中包含 DMSO。

12. 用如下方法计算测试样品中探针的量：

   (ng 对照中的量×测试行的稀释因子)/10＝探针浓度（ng/μl）

   也就是说，利用表 8-3，如果探针稀释 8 倍的强度与标准稀释 10 倍的强度相等，浓度则计算为 (0.19 ng×1280)/10＝24 ng/μl.

627

13. 未使用的探针（来自步骤 5）保存在－70 ℃。

   反复冰冻/融化探针会降低保存时间。为克服这个问题，可以根据预期的使用，将溶液分成小的分量分装存放。在－70 ℃下保存时，探针可以保持完全活性至少 18 个月。

   使用中向组织阵列上加入 100 μl 浓度为 200 ng/ml（推荐值）的探针/杂交液时，200 ng 探针足够最多 9 片玻片使用。

14. 通过琼脂糖凝胶电泳确定标记 cRNA 转录产物的长度和产量。向与 DIG 标记 RNA 相邻的泳道加 1 μg 线性 DNA 模板。加入 RNA 和 DNA 分子质量标准。用溴化乙啶染色凝胶，观察模板和标记的核酸。

组织制备和杂交

15. 按照实验方案 1 所述固定，包埋组织和排列阵列。
16. 将厚度为 5 μm 的组织阵列切片裱贴在 Superfrost Plus 载玻片上，并在 60 ℃烘烤 2 h。

   使用经 DEPC 处理的水漂浮切片，并转移到载玻片上。

   存放于 4 ℃干燥器中预先制备的切片也可以使用。

为了减小杂交前和杂交过程中标记探针的降解，参看本书第三部分中关于 RNA 操作特别注意事项里概括的预防措施。

17. 在室温下让切片通过下面一系列新鲜配制的溶液进行脱蜡和重新水化。在切片转移到下一种溶液中之前要完全沥干。

Histolene	10 min
Histolene	10 min
无水乙醇	5 min
90% 乙醇	5 min
70% 乙醇	5 min

18. 将载玻片在 0.2 mol/L HCl 中振荡洗涤 20 min。
19. 对于非放大式 ISH，下面直接进行步骤 20。对于放大式 ISH，将载玻片在 3% H_2O_2 中振荡洗涤 10 min，以灭活任何内源的过氧化物酶活性。室温下将载玻片在 PBS 中轻微搅动，漂洗 5 min。
20. 将载玻片在经 DEPC 处理过的灭菌水中漂洗 2 次，每次 5 min，并轻微搅动。
21. 用蛋白酶 K 消化阵列以暴露靶 mRNA，通常酶的最终浓度为 1 μg/ml。将载玻片在预热过（37℃）的蛋白酶 K 缓冲液中平衡 5 min。加入适量的蛋白酶 K，通过数次反复浸入和取出载玻片来混合溶液。在 37℃下孵育载玻片 30 min，不要振荡。

    参见本方案介绍部分的“暴露靶 mRNA”和表 8-2 里建议的优化方案。

22. 将载玻片浸入 0.2%甘氨酸/PBS（预冷至 4℃）中 10 min，并轻微振荡，终止蛋白质的消化。
23. 将载玻片浸入 0.1 mol/L 三乙醇胺中，并迅速加入 100%乙酸酐至最终浓度为 0.25%，以乙酰化组织阵列。取出和重新浸入玻片数次，以确保有效的乙酰化。室温下将玻片振荡孵育 5 min。

    乙酸酐会迅速分解，因此必须从制造商提供的容器中直接加入；0.1 mol/L 的三乙醇胺要在使用前新鲜配制。在制备 0.1 mol/L 溶液时确保精确地量取高黏度的三乙醇胺。

    注意：三乙醇胺对皮肤和黏膜有刺激性 。该步骤要在通气良好的通风橱中进行。

24. 将载玻片在经 DEPC 处理的灭菌水中振荡漂洗两次，每次 5 min。
25. 将 40 μl 核糖核酸溶液加入 1 ml 预杂交缓冲液中，然后将 100 μl 溶液点样到盖玻片上。轻轻放下一片带有组织阵列的玻片到预杂交缓冲液液滴的表面，直到盖玻片黏附到载玻片上。在合适的杂交温度下，将玻片置于用大量 40%甲酰胺湿润的杂交盒中预杂交 2 h。

    请参看本部分介绍中的讨论“温度条件”来确定杂交温度。

26. 将 200 ng（放大式 ISH 用 50 ng）DIG 标记的探针用经 DEPC 处理的水稀释到 80 μl。加入 40 μl 核糖核酸溶液，并将混合物在 65 ℃ 孵育 5 min 以变性标记的 RNA。立即将试管转移到冰上。

    请参看本实验方案介绍中关于探针浓度的讨论和表 8-2。

探针的制备和杂交

27. 向 1 ml 预杂交缓冲液中加入：

1 μl　　100 mmol/L 二硫苏糖醇

10 μl　　10% 硫代硫酸钠

10 μl　　10% SDS

将封闭的试管用力快速翻转数次以混合溶液。任何时刻都要保持杂交缓冲液处于冰块的冷却之下。

28. 将步骤 26 制备的探针，与步骤 27 中制备的 880 μl 杂交液合并。充分混合溶液并迅速转移到冰上。

涡旋振荡通常不足以充分混合高黏度的溶液。将试管向塑料支架的横档快速而有力的碰撞。溶液中出现的气泡确保了发生充分混合。将溶液在微型离心机中，4 ℃下以最高速度离心除去这些气泡。

1 ml 杂交缓冲液/探针混合物足够用于在 9 片玻片上进行 ISH。

29. 用镊子把盖玻片从预杂交的组织阵列上轻轻地滑开。将玻片边缘在灭过菌的吸水纸巾上轻敲 2 或 3 次除去过量的预杂交缓冲液。不要试图除去全部的预杂交缓冲液。加入 100 μl 没有气泡的探针溶液（步骤 28）和一片清洁的盖玻片。在密闭的、用大量 40%甲酰胺湿润的杂交盒中，将玻片在合适的杂交温度下温育 16 h。

为了减少组织从玻片上缺失，请参看实验方案 3 最后部分中表 8-6 里的讨论。

气泡会造成假阴性。施压于盖玻片来驱赶气泡，会使得载玻片和盖玻片之间可供探针扩散的空间减少，从而带来不想要的结果。控制气泡最佳的方法是从开始就避免产生气泡，这需要仔细地将探针加到盖玻片上，然后再将载玻片放到盖玻片上。

高黏度的杂交缓冲液/探针溶液会扩展到它们所接触的全部玻璃表面。要确保包含不同探针的全部玻片，特别是阴性对照，在杂交盒内相互之间分隔良好。

为了避免虚假的染色，要确保所有的盖玻片都没有任何灰尘。

缩短从杂交盒/箱中移开玻片和在杂交盒/箱中再次温育之间的时间间隔，以保持玻片的温度，从而保证杂交反应的严格程度。

洗涤和探针的观察

30. 在 2×SSC 中搅动 5 min，轻轻地从载玻片上除去盖玻片。在振荡水浴中进行下面一系列的洗涤：

SSC 浓度	时　间	温　度
2×SSC	15 min	室温
2×SSC	15 min	杂交温度
1×SSC	15 min	杂交温度
0.1×SSC	15 min	杂交温度

如果盖玻片在自身的重力作用下与载玻片分离，洗涤前载玻片不必在 2×SSC 中搅动。

31. （可选的）如果需要用 RNase A 处理，37 ℃下将载玻片在 RNase 消化缓冲液中振荡洗涤 3 次，每次 5 min。或者前进到步骤 33 时，将 RNase A 加入到 RNase 消化缓冲液中，直至最终浓度为 20 μg/ml，37 ℃下将玻片振荡孵育

60 min。

只有当背景染色高得无法接受时才使用 RNase A 消化。见本实验方案介绍中关于 RNase A 消化的部分。

32. 将载玻片在预冷至 4 ℃的 PBS 中搅动 2 次，每次 5 min，以终止 RNase A 消化。

33. 将载玻片在顺丁烯二酸缓冲液中搅动 5 min。

34. 向载玻片上加入 100 μl 封闭溶液，以封闭抗 DIG 抗体的非特异性结合。在湿润的盒中进行该步骤，或者使用盖玻片防止封闭溶液的蒸发。室温下将载玻片孵育 30 min。

当要处理的载玻片数量超过 20 片时，为方便本步骤操作，可以在载玻片盒里，将载玻片全部浸入合适体积（如 150 ml）的封闭溶液中。

35. 对于非放大式 ISH，向每片载玻片上加入 100 μl 在封闭溶液中稀释到 1∶1000 的抗 DIG-AP 抗体。将载玻片在室温下孵育 60 min。对于放大式 ISH，向每片载玻片上加入 100 μl 在封闭溶液中稀释到 1∶100 的抗 DIG-AP 抗体，并将载玻片在室温下孵育 60 min。

关于抗 DIG-HRP 抗体的建议浓度，参考表 8-2。

630

36. 室温下将玻片在顺丁烯二酸缓冲液（pH 7.5）中振荡洗涤 3 次，每次 10 min。

对于非放大式 ISH：前进到步骤 37。

仅适用于放大式 ISH：

a. 向每片载玻片上加入 100 μl DNP 放大底物。室温下，将玻片在用蒸馏水湿润的密闭盒中孵育 10 min。

关于推荐的孵育时间，参考表 8-2。

b. 室温下将玻片在顺丁烯二酸缓冲液（pH 7.5）中振荡洗涤 3 次，每次 10 min。

c. 将 100 μl 用封闭溶液按 1∶50 稀释后的抗 DNP-AP 抗体，加入到每片载玻片上。将玻片在室温下孵育 30 min。

d. 室温下将玻片在顺丁烯二酸缓冲液（pH 7.5）中振荡洗涤 3 次，每次 10 min。

关于推荐的抗 DNP-AP 抗体浓度，参考表 8-2。

37. 仔细擦干载玻片，不要碰到组织切片。用防水记号笔给切片画圈。

38. 加入足量新鲜配制的 NBT/BCIP 显色试剂，直至覆盖组织阵列。将玻片在黑暗中孵育。

用显微镜查看显色情况，直到获得期望的染色强度（通常 1～24 h）。每 6 h 更换一次新鲜的 NBT/BCIP 显色物质。

39. 将玻片在蒸馏水中漂洗，用水溶性染料复染（见下面注解），并用合适的封固剂盖上盖玻片。

注意：乙醇会溶解和萃取显色反应的产物。在复染和封片的时候，避免已经显色的玻片与 histolene 和乙醇接触。

Kaisers 甘油凝胶，Histomount 或 Clearmount 等水溶性封固剂的使用都很方便。

个人偏好决定了对探针信号的对比（深紫色/蓝色）和显示组织形态学的复染方法的选择。推荐使用甲基绿、细胞核坚牢红或曙红。将玻片浸入复染溶液，直至得到希望的复染程度。

表 8-4 原位杂交疑难解答指南

表 现	鉴 别	可能的解决方案
信号弱，或者全无	杂交严格程度太高	使用较低杂交温度
	组织阵列被蛋白酶过度消化（导致 mRNA 从切片流失）	重新优化蛋白酶 K 的消化
	蛋白酶 K 对组织阵列消化不足（未能有效暴露 RNA）	重新优化蛋白酶 K 的消化
细胞核信号	受到污染或者是试剂过期	准备一套新鲜配制的溶液
使用 SNCP 或者正义探针时有强烈的非特异性信号	通常是由杂交严格程度低造成	增加杂交严格程度；杂交后用 RNase A 处理标本
杂交过程中或者杂交后处理时组织缺失	组织与载玻片结合不牢固	组织脱蜡前至少烘烤 2h。如果问题依然存在，尝试再次烘烤组织两次（每次 2h）。两次烘烤之间让载玻片冷却到室温。使用高于组织表面的盖玻片，如 Lifter Slips。
阵列内部变异	组织阵列单元采用了不同的固定方法	用无尘的塑料薄片代替盖玻片。选用由同种方法固定的组织制备阵列，或者为每种组织和固定方法优化 ISH 参数（见表 8-2）

方案 3

组织阵列上的免疫组织化学

Melanie Trivett（Peter MacCallum 癌症中心，东墨尔本，维多利亚 3002，澳大利亚）

免疫组织化学（immunohistochemistry，IHC）通过抗原-抗体反应识别细胞或组织抗原（Miller 1996）。通过直接标记一抗，或者二次标记结合的一抗来检测抗体的结合位点。这里介绍的免疫组织化学方法由三个步骤组成，使用了标记的链霉亲和素-生物素（streptavidin-biotin，LSAB），组织切片首先是和与感兴趣抗原直接相抗的一抗孵育，然后与和一抗可发生结合的生物素化连接试剂（二抗）孵育。最后，通过辣根过氧化物酶-链霉亲和素缀合物显示生物素化抗体的位置。过氧化物酶的活性通过有过氧化氢存在下，可以在靶抗原位置产生不溶性沉淀的显色物质来观察（Miller 1996）。关于链霉亲和素和生物素的相互作用，以及 HRP 催化的反应见 Sambrook and Russell 2001，A9.35～A9.37 和 A9.45 页。复染切片，就可以通过光学显微镜观察到与抗体位置形成对照的组织形态学。通过应用组织切片常规分析中所建立的方法，IHC 已经成功地应用于组织阵列。

组织的固定

Hopwood（1996）对组织学中的固定和固定剂给出了极佳的讨论。理想情况下，固

定要尽可能地保持组织的自然状态。尽管没有一种固定剂对所有抗原都是最佳的，10%中性（pH 7.5）缓冲的福尔马林（NBF）由于应用方便、可适用于广泛类型的组织、和能够可靠保留细胞形态的特点而在组织病理学中得到广泛的应用（Hopwood 1996；Miller et al. 2000）。然而，在使用保存的组织块制备组织阵列时，研究者无法把握所应用的固定剂或者固定的时间（Moch et al. 2000）。

福尔马林固定方法的一个潜在的问题是隐蔽组织的表位。表位是抗原的一部分，与一抗发生反应。隐蔽可能是由于固定剂与和表位相邻的氨基酸发生反应导致构象的变化，或者与表位自身发生反应。在绝大多数情况下，抗原显露可以用蛋白水解酶处理实现，或者通过对福尔马林固定和石蜡包埋（formalin-fixed paraffin-embedded，FFPE）的组织切片进行抗原恢复（Miller 1996；Taylor et al. 1996；Miller et al. 2000）。许多商业化的抗体已经过 FFPE 切片检验，并在抗体说明书中提供了一些抗原暴露的建议方法。

633

蛋白质水解消化

许多蛋白质水解酶，包括胰蛋白酶，胃蛋白酶和蛋白酶 K 都可以用于在组织切片中显露抗原（Miller 1996）。酶促消化的成功与否受到福尔马林固定时间长短的影响（Miller 1996），并依赖于寻找合适的条件以避免过度或者不完全消化。虽然很多种抗原可以通过抗原恢复去隐蔽（见下文），但是对一些抗原来讲，蛋白水解消化仍然是最佳的方法。最佳的去隐蔽和恢复方法必须由各个实验室针对每种抗原来确定（Miller 1996）。

抗原的恢复

抗原恢复过程，也被称为热诱导的表位恢复（heat-induced epitope retrieval，HIER），通过高温加热来暴露福尔马林固定组织切片的抗原，从而提高大量抗体的免疫组化染色灵敏度，并降低背景染色（Taylor et al. 1996；Shi et al. 1997；Miller et al. 2000）。可使用的不同加热方式有微波、压力锅、置于微波中的压力锅、蔬菜蒸笼、高压灭菌锅和水浴。最佳结果与温度和加热时间的效果相关，而与加热方式无关（Taylor et al. 1996；Miller et al. 2000；Shi et al. 1997）。抗原恢复可让多种经福尔马林固定时间长短各异的石蜡组织切片的免疫反应活性变得一致（Shi et al. 1997）。然而，某个特定表位去隐蔽的效率受到恢复溶液 pH 和化学组成的影响。因此，在每个实验室中，针对每种抗体最佳的抗原恢复条件必须根据经验来确定（Shi et al. 1995；Pileri et al. 1997）。

用于 IHC 切片的保存

石蜡切片在长期保存中，由于抗原的丢失或者隐蔽，免疫染色的强度和程度都会发生衰减（Wester et al. 2000）。抗原性降低的机制尚不清楚。不同抗原间的抗原性丢失程度是各不相同的，看起来似乎受到温度的影响（Wester et al. 2000）。当进行切片和保存用于 IHC 的组织阵列切片时，必须要考虑到这些问题。当新的抗体应用于组织阵列进行优化时，对新制备和保存的切片进行比较是有帮助的。如果可能，尽量避免使用

固定方法、保存时间和温度不相同的切片（Wester et al. 2000）。

优化 IHC 方案

一抗

IHC 中使用的抗体有两种类型：单克隆或多克隆。单克隆抗体识别位于靶抗原上的单个表位，而多克隆抗体几乎总是识别多个表位（Miller 1996）。

单克隆抗体的一个优点是它为已知特异性的单纯试剂提供了一个稳定的来源（Miller 1996）。然而，多克隆的抗血清受隐蔽作用的影响较小，这是因为它可以识别多个表位，而这些表位被固定作用隐蔽或者经去暴露处理而显示的能力也各不相同（McNicol and Richmond 1998）。 634

对照

已知包含要研究抗原的组织（阳性组织对照）用做抗体/IHC 优化中消除潜在的假阴性结果。已知不产生所要研究抗原的组织（阴性组织对照）用于确证免疫组化染色的特异性。一旦抗体得到优化，其他对照，如非免疫血清的替代（用于多克隆抗血清）或者同型匹配的免疫球蛋白（用于单克隆抗体），就可以应用于免疫组织化学染色的验证（Miller 1996）。

抗体滴定

对新的一抗进行测试时，要进行一系列较宽范围的稀释，以确保假阴性结果不会发生（Miller 1996）。如果一抗以前曾在免疫组织化学中应用过，通常可以在文献中找到抗体稀释的初始值。商品化抗体的说明书中也可能会建议合适的稀释范围。

抗原去隐蔽

抗原去隐蔽的效率随抗原的不同而变化，并受到固定方法的影响（Shi et al. 1995；Taylor et al. 1996；McNichol and Richmond 1998；Frost et al. 2000；Miller et al. 2000）。商品化抗体所提供的说明书通常给出了抗原去隐蔽方法。如果一抗先前没有在 IHC 中使用过，则需要尝试一系列的去隐蔽实验方案（包括不做处理，酶消化和抗原恢复）。对同样的抗体，不同固定剂产生的结果可能不同，但在组织阵列中，如果阵列从保存的石蜡块制备，研究人员可能无法控制固定的方式。

推荐一组实验方法（见表 8-5；也见 Shi et al. 1995；Taylor et al. 1996；McNichol and Richmond 1998）来为任何新的一抗建立最佳的抗原恢复条件。我们建议保留一片玻片不做处理，用作对照。

作者所在实验室常规用柠檬酸缓冲液（pH 6.0）在压力锅中处理 2 min 进行抗原恢复，并且大量抗体都得到了好的免疫组化染色。不同成分和 pH 的抗原恢复缓冲液可以购买得到。然而，最佳的抗原恢复条件不仅随特定抗原而变化，而且不同实验室中分析同种抗原时也是如此。因此，抗原恢复要看作是一种魔法和烹调法的混合物。

表 8-5 通用的"实验组"用于发展最佳试验方案

柠檬酸缓冲液	pH 1.0～2.0	pH 6.0	pH 10.0～11.0
Tris-HCl 缓冲液	pH 1.0～2.0	pH 7.0～8.0	pH 10.0～11.0
高压灭菌/压力锅[a] 120 ℃下 10 min 微波	玻片 1	玻片 2	玻片 3
100 ℃下 10 min 微波或水浴	玻片 4	玻片 5	玻片 6
90 ℃下 10 min	玻片 7	玻片 8	玻片 9

经许可，从 McNicol 和 Richmond (1998) 复制。

635

a 如果没有，微波处理 20 min。

检测系统

目前已有现成的，用于 IHC 的试剂盒。我们实验室常用 LSAB 2 试剂盒（DAKO）或者更为灵敏的 LSAB ＋试剂盒（DAKO），作为过氧化物酶介导的对人体组织一抗的检测。这些试剂盒，作为现成的二抗和标记，检测单克隆和多克隆抗血清。也有碱性磷酸酶检测系统的试剂盒。许多商品化的试剂盒由于与试剂盒内提供的二抗混合物之间有交叉反应，可能不适用于非人类组织。有许多商品化的试剂盒可以用于小鼠组织上小鼠单克隆抗体的检测。

一些抗体可能需要更加灵敏的检测系统，如通过 biotinylated tyramine 放大信号（King et al. 1997）。免疫组化的信号放大试剂盒可以购买得到。

材料

注意：带〈!〉材料正确的操作方法见附录 3。

重要事项：除非特别说明，试剂在室温下保存。

缓冲液和溶液

关于储存液、缓冲液和试剂的成分见附录 2。

将储存液稀释至合适的浓度。

氯化钙（1%）〈!〉

氯化钙（二水合物，m. w. 147.02）	5 g
蒸馏水	500 ml

将水和氯化钙混合。在 4 ℃保存，用作储存液。

氯化钙（0.1%）

1%氯化钙（见上）	50 ml
蒸馏水	450 ml

由上述储存液制备。在 4 ℃保存，酶消化时根据需要使用。

乙醇（70% 和无水乙醇）

胎牛血清（fetal bovine serum，FBS）（10% *V/V*）

FBS	1 ml
50 mmol/L Tris-HCl（pH 7.6）/0.05% Tween-20	9 ml

在－70 ℃按 1 ml 每份保存 FBS。每次免疫组织化学实验配制新鲜的 10% FBS。任何残留的 10% FBS 可在 4 ℃保存，留待次日使用。

过氧化氢（3%）〈!〉

30%过氧化氢（100 体积）	5 ml
蒸馏水	45 ml

3%过氧化氢在使用前新鲜配制。

速溶苏木精（可从 Shandon 得到）

636

速溶苏木精 2× 成分 A

速溶苏木精 2×成分 B

蒸馏水 2 L

戴上手套。将速溶苏木精成分 A 和 B 溶解在 2 L 蒸馏水中。充分混合溶液，并在黑暗中保存。使用前过滤溶液。经过滤的苏木精可在染色皿中放置约两周时间。

蛋白封闭试剂

蒸馏水	25 ml
蛋白封闭剂（2×）	25 ml

蛋白封闭剂（2×）可从 Lipshaw Immunon 得到。4 ℃保存。

Scotts 自来水

碳酸氢钠（m. w. 84.01）	7 g
$MgSO_4 \cdot 7H_2O$（m. w. 246.48）	20 g
自来水	定容到 2000 ml

将各组分溶解于自来水。

柠檬酸钠（1 mol/L，pH 6.0）

柠檬酸，三钠盐（m. w. 294.1）〈!〉	294.1 g
蒸馏水	定容到 1000 ml

将柠檬酸（三钠盐）溶解于水中，定容至 1000 ml。用浓盐酸〈!〉. 将 pH 调整到 6.0。存放于 4 ℃作为储存液。

柠檬酸钠（10 mmol/L，pH 6.0）

1 mol/L 柠檬酸钠（pH 6.0）（见上）	20 ml
蒸馏水	1980 ml

对于每次需要抗原恢复的免疫组织化学，从上述储存液新鲜配制。

氢氧化钠（1 mol/L）〈!〉

氢氧化钠（m. w. 40）	4 g
蒸馏水	定容到 100 ml

用作储存液，保存在密闭的塑料容器中。

氢氧化钠（50 mmol/L）

1 mol/L 氢氧化钠（见上）	5 ml
蒸馏水	95 ml

该浓度用于调节组织切片蛋白水解消化的胰蛋白酶/氯化钙溶液的 pH。由上面的氢氧化钠 1mol/L 储存液制备。

Tris-HCl（1mol/L，pH 7.6）/1% Tween-20（单月桂酸酯）

Tris 碱（m. w. 121. 14）	242.28 g
蒸馏水	1600 ml
浓盐酸	120 ml
Tween-20	20 ml

这里用作储存液。将 Tris 碱溶于 1600 ml 蒸馏水。搅拌直到溶解，然后加入 20 ml Tween-20。小心加入 120 ml 浓盐酸，让缓冲液冷却至室温。用浓盐酸调节 pH 到 7.6。在室温下让 pH 稳定过夜，次日早晨根据需要调节 pH。用蒸馏水定容到 2000 ml。重新检查 pH，如果需要再行调节。将缓冲液分装成 500 ml 等份，在 4 ℃保存。

Tris-HCl（50mmol，pH 7.6）/0.05% Tween-20

1 mol/L Tris-HCl/1% Tween-20（见上）	500 ml
蒸馏水	9500 ml

637

这是免疫组织化学染色的工作缓冲液，从上面的 1 mol/L Tris-HCl 储存液配制。

自来水/0.05% Tween-20

Tween-20	5 ml
自来水	10 L

二甲苯或者二甲苯替代物，如 histolene〈!〉

酶和缓冲液

胰蛋白酶（10%）〈!〉

胰蛋白酶（IIs 型）	10 g
50 mmol/L Tris-HCl（pH 7.6）/0.05% Tween-20	100 ml

胰蛋白酶可以从 Sigma（T 7409）获得。在化学通风橱中配制 10%的胰蛋白酶。向胰蛋白酶容器中加入 50 mmol/L Tris-HCl（pH 7.6）/0.05% Tween-20，混合以溶解胰蛋白酶。将溶液倒入容积为 100 ml 的 Schott 瓶中，并将剩下的 50 mmol/L Tris-HCl（pH 7.6）/0.05% Tween-20 加入并混合。在微量离心管中将溶液分成 0.5 ml 等份。在－20 ℃保存。

细胞和组织

组织阵列块

按实验方案 1 制备。

专用设备

抗原去隐蔽

- 玻璃 Coplin 广口瓶
- pH 计
- 不锈钢压力锅和电热板
- 水浴

切片

- 干热鼓风烤箱
- 浮选水浴

玻璃载玻片（Superfrost Plus，Fisher Scientific）

显微镜用薄片切片机

玻片加温用热板

免疫组化染色

AEC＋（DAKO）或 DAB＋（DAKO）显色物质〈!〉

多种显色物质可用于过氧化物酶的标记。使用 3-氨基-9-乙基咔唑（3-amino-9-ethylcarbozole，AEC）会得到一种红色的终产物，而二氨基联苯胺（3′,3′-diaminobenzidene，DAB）则得到棕色产物。AEC 的产物可溶于酒精，因此在染色完成后必须使用水溶性封固剂封片。预先稀释的这些和其他显色物质可以购买得到（Miller 1996）。

LSAB 2（DAKO）和 LSAB＋（DAKO）检测试剂盒

玻片孵育盒（如玻璃或者有机玻璃盒）

脱蜡用染色皿（5）

洗涤用染色皿

染色架（耐热的）

封片

水溶性封固剂（来自 Biomedia 的 Crystal Mount，或者来自 DAKO 的 Ultramount）

玻璃盖玻片

含树脂的封固剂（DPX）

用于脱水的染色皿（5）

方法

638

本方法适用于石蜡包埋的组织阵列。除非另有说明，全部操作步骤在室温下进行。最佳的条件（去隐蔽，一抗的滴定，检测试剂盒）必须由每个实验室确定。

组织切片的制备

1. 用显微镜用薄片切片机对组织阵列进行 3 μm 切片。

   切取足量的切片以试验每种去隐蔽的方法。对每种将要尝试的去隐蔽方法，记住在每个组织阵列块中包括一个阴性对照。

2. 将切片漂浮于充满蒸馏水的浮选水浴中。

   将水浴的温度设置为低于切片石蜡熔点 5 ℃。

3. 将切片放置在带电荷的载玻片上，如 Superfrost Plus，以增强组织的附着。

   在浮选水浴中加入组织黏合剂可能会影响组织切片与带电荷载玻片的结合。

   在浮选水浴中使用自来水可能会影响组织切片和载玻片的结合。在组织切片与在玻片结合不受影响的情况下，可以用自来水代替浮选水浴中的蒸馏水。

   如果 IHC 中使用自动染色仪染色，要保证切片放置在载玻片上的位置适合于染色试剂的分配。

4. 将载玻片在热台上（温度设为约 45 ℃）烘干，至少 30 min。

   在继续下面操作之前，确保组织阵列切片下方所有水分全部蒸发。

5. 在 56～60 ℃下，将载玻片在鼓风干燥箱中烘烤 30 min。

如果组织与在玻片的附着较差，载玻片可以烘烤最多 2 h。

6. 将切片叠在一起保存在载玻片盒中，以隔离空气和水分。在最上层切片的上方覆盖一片空白的载玻片，用橡皮筋将切片扎在一起，并将切片于室温下在黑暗中保存。

对一些抗原，将切片保存在 4℃下（Wester et al. 2000）可能减少抗原性的损失。

7. 在载玻片上记录下日期以解释与切片保存相关的问题。

切片的脱蜡和重新水化

在开始免疫组织化学染色之前，要确保切片石蜡从组织阵列中溶解，并且切片被重新水化。切片可以在干热鼓风干燥箱中短暂加热后进行迅速脱蜡（当石蜡熔化时切片脱蜡较快）。脱蜡的系列操作包括两次更换二甲苯（或者二甲苯替代物），两次更换无水乙醇和一次更换 70％的乙醇。使用防热的支架，这样切片可以放入热的缓冲液中进行抗原恢复（如果需要）。一旦脱蜡和重新水化后，切片不能变干燥，这一点很重要。脱蜡要在通风橱中于室温下进行。

639

8. 将切片放入二甲苯，或者二甲苯替代物，如 histolene 中，时间 5 min。
9. 将切片转移到下一组二甲苯/histolene 中，时间 5min。
10. 将切片转移到无水乙醇中，时间 60 s。
11. 将切片转移到下一个装有无水乙醇的器皿中，时间 60 s。
12. 将切片转移到 70％乙醇中，时间 60 s。
13. 将切片转移到自来水中。

为了保证结果的一致性，脱蜡系列试剂要定期更换，例如，如果使用容量 200 ml 的染色皿，每 48 片切片更换一次。在进行切片脱蜡时，准备抗原去隐蔽所需的溶液。如果通过胰蛋白酶消化对抗原去隐蔽，要预热缓冲液；对抗原去隐蔽缓冲液进行加热。

抗原去隐蔽的方法

14. 继续进行抗原去隐蔽（如果需要）。

关于抗原去隐蔽和酶消化的方法，请见本方案结束部分关于替代实验方案：抗原去隐蔽和胰蛋白酶消化的信息栏。阴性对照必须和实验组切片接受同样的处理。

免疫组织化学染色

本方法借助于一种生物素化的二抗和链霉亲和素，利用辣根过氧化物酶进行检测。在开始免疫组织化学染色步骤时，切片要用过氧化氢处理，以阻遏切片中内源的过氧化物酶活性。根据制造商的说明，我们实验室通常使用 LSAB 2（DAKO）和 LSAB＋（DAKO）试剂盒检测。除非特别说明，所有操作步骤在室温下进行。

15. 将切片放入装有新鲜配制的 3％（V/V）过氧化氢中 10 min（以阻遏内源性的过氧化物酶）。
16. 将切片在流动的自来水中洗涤 3 min。
17. 将切片置于装有蛋白质封闭溶液的器皿中 20 min。

将载玻片平放在湿润的孵育盒中，向每张切片加入蛋白质封闭试剂，可以减少蛋白质封闭试剂的用量。关键要使得试剂覆盖整张切片的表面。

一些抗体可能不需要蛋白质封闭步骤。这要根据经验来确定。

18. 在进行蛋白质封闭步骤的时候，配制一抗和适合的阴性对照试剂。用 10% FBS/50 mmol/L Tris-HCl（pH 7.6)/0.05% Tween-20 稀释所需用量的一抗。

    阴性对照试剂的制备：对于鼠单克隆抗体，免疫球蛋白对照要和一抗稀释后具有相同的最终免疫球蛋白浓度。多克隆抗体的情形与此类似，非免疫血清的最终蛋白质浓度要和一抗稀释后的蛋白质浓度相同。

640

19. 将蛋白质封闭溶液从切片倾倒掉。不要漂洗切片。
20. 用吸水纸巾擦干切片周围，以除去过量的封闭试剂。
21. 将载玻片放入湿润的孵育盒，并向每张载玻片加入经适当稀释的一抗血清，要确保切片阵列上所有的点都被覆盖。向阴性对照加入阴性对照试剂。
22. 在室温下将载玻片孵育 30 min。

    要得到最佳结果，可能需要在室温下将一抗孵育 1～2 h，或在 4 ℃过夜。这要根据经验来确定。

23. 用流动的 50 mmol/L Tris-HCl（pH 7.6)/0.05% Tween-20 从载玻片上漂洗去一抗。
24. 将载玻片放入 50 mmol/L Tris-HCl（pH 7.6)/0.05% Tween-20 浴中，洗涤切片 5 min。

    洗涤时最好在振动台上轻摇。如果无法摇动，将洗涤时间延长到 10 min。

25. 沥干载玻片，并用吸水纸巾擦干切片周围多余的缓冲液。
26. 加入二抗，并将载玻片放置湿润的孵育盒中。
27. 按照二抗制造商的说明，对载玻片进行孵育。
28. 用流动的 50 mmol/L Tris-HCl（pH 7.6） 10.05% Tween-20 洗涤载玻片。
29. 将载玻片置于架上，在 50 mmol/L Tris-HCl（pH 7.6)/0.05% Tween-20 浴中洗涤 5 min。

    如果需要的话，这个阶段中切片可以在缓冲液中停留最多 30 min。

30. 沥干载玻片，并用吸水纸巾擦干切片周围。
31. 加入链霉亲和素过氧化物酶，并将载玻片放在湿润的孵育盒中。
32. 按照制造商的说明孵育载玻片。
33. 用流动的 50 mmol/L Tris-HCl（pH 7.6)/0.05% Tween-20 漂洗载玻片，并将载玻片放在载玻片架上。
34. 用 50 mmol/L Tris-HCl（pH 7.6)/0.05% Tween-20 洗涤载玻片 5 min。

终产物反应

35. 用吸水纸巾擦干切片周围除去过量的缓冲液。
36. 向每片载玻片加入显色物质（AEC 或者 DAB)。

    注意：阅读制造商关于显色物质的安全说明，因为大多数显色物质可能会致癌，因此操作时要注意。

37. 按照制造商的说明孵育载玻片（通常为5～10 min）。

38. 从载玻片上倾倒去显色物质。

将显色物质收集到废物缸中，并按照当地的规章处理。

39. 在流动的自来水中洗涤载玻片。

复染

40. 在切片复染前，用吸水纸巾除去苏木精溶液表面任何堵塞物质。

41. 将载玻片放入苏木精中约5 s，对细胞核染色。

641

细胞核染色需要的时间取决于苏木精的新鲜程度。理想情况下，复染要得到蓝色的细胞核，而细胞质则几乎很少或没有染色。

42. 从苏木精中取出载玻片，在流动的自来水中洗涤。

43. 将切片放入Scott's Tap Water 30 s，使苏木精变蓝。

44. 在流动的自来水中洗涤载玻片。

45. 将载玻片转移到盛有0.05% Tween-20/自来水的器皿中。

水封片和盖片

46. 对载玻片处理以封片。根据免疫组化染色时使用显色物质的不同，采用下列方法之一。

溶解于乙醇的显色物质：如果使用了溶于乙醇的显色物质，如AEC，在用树脂封固剂，如DPX和二甲苯（或者二甲苯替代物）将盖玻片放到载玻片上之前，载玻片必须用水介质封片。如果在染色过程中混合使用了几种显色物质，那么不溶于乙醇的显色物质也可以用这种方法进行处理。

a. 一次沥干数片载玻片，并用吸水纸巾擦干切片周围除去过多的液体。

b. 加入足量的水性封固剂以覆盖切片。

c. 水平拿住每张载玻片，并仔细地将封固剂扩展到载玻片上。

d. 除去载玻片上任何气泡。

e. 将载玻片放在60 ℃的热台上，以固化封固剂。

f. 将载玻片放到盖玻片上之前，冷却载玻片至室温。

g. 使用DPX和二甲苯，或二甲苯替代物，将盖玻片放到载玻片上。

h. 除去盖玻片下的气泡。

不溶于乙醇的显色物质：对不溶于乙醇的显色物质，无需使用水性封固剂，如DAB。相反，将盖玻片放到载玻片上之前，要除去切片的水分。该方法不适用于溶于乙醇的显色物质。

a. 将载玻片放入70%乙醇中2 min。

b. 将载玻片转移入无水乙醇中30 s。

c. 将载玻片转移入新鲜的无水乙醇中30 s。

d. 将载玻片转移入二甲苯（或二甲苯替代物）中2 min。

e. 将载玻片转移入新鲜的二甲苯。

上述操作步骤中要使用单独的脱水试剂组，以避免污染脱蜡试剂组。

f. 用DPX将玻璃盖玻片放到载玻片上。

表 8-6 免疫组织化学的疑难解答指南

问　题	可能的解决方法
抗原恢复步骤中丢失组织切片	• 使用有涂层的载玻片。 • 不要向水浴中加入黏合剂。 • 在水浴中使用蒸馏水。 • 在干热烘箱中延长载玻片的烘烤时间。
在阳性对照中没有观察到免疫反应性	• 检查用于染色的阳性对照组织。 • 检查所使用全部缓冲液的 pH。 • 检查一抗，二抗和标记按照正确的次序使用。 • 重新考虑一抗的稀释。 • 延长一抗的孵育时间。 • 重新考虑抗原恢复的实验方案（尝试一种不同的抗原恢复缓冲液，或者用不同的酶消化）。 • 尝试更灵敏的检测系统；一些抗原仅能为放大系统，如生物素化酪胺（biotinylated tyramine）检测（King et al. 1997）。有商品化试剂盒供应。 • 切取新鲜切片以避免失败的可能性。
IHC 批次间染色的改变	• 检查载玻片的保存条件。 • 检查对比新鲜和保存的切片，以确定是否发生抗原性的丢失。 • 检查抗原恢复方案的操作前后一致。 • 抗原恢复后（如果使用）保持冷却时间一致，如 20 min。 • 检查酶消化缓冲液和抗原恢复缓冲液的 pH。
阴性对照被染色	• 检查过氧化物酶封闭步骤。使用新鲜的 30% 过氧化氢配制 3% 溶液。 • 按照如下方法检查内源生物素活性（endogenous biotin activity，EBA）：将一张切片脱水，去隐蔽抗原（如果使用了抗原去隐蔽操作），封闭内源过氧化物酶活性，然后加入链霉亲和素标记，再加入显色物质。切片显色说明 EBA。可以得到商品化的试剂盒用于封闭 EBA，并且在组织学或者 IHC 讨论页上有许多自制的配方。在一些组织中，EBA 随着抗原恢复而增强（Naish 1989）。在进行任何抗原去隐蔽/恢复操作后，进行 EBA 封闭步骤。 • 检查所使用全部缓冲液的 pH，如果需要的话，检查所使用全部 pH 计的准确度。
切片的非特异性免疫染色	• 过度染色，重新滴定一抗（染色要随着稀释因子的增加而降低强度）。 • 尝试增加洗涤时间和不同抗体（一抗和二抗）孵育之间的清洗次数，以除去未结合的抗体。 • 尝试不同的抗体稀释物（可以得到旨在减低背景染色的商品化配置品）。 • 如果没有使用的话，加入蛋白质封闭步骤。注意封闭的效果可能随使用的封闭剂不同而变化。 • 一些去隐蔽方案可能会产生虚假染色，如背景核染色（Miller 1996；Taylor et al. 1996；Miller et al. 2000）。尝试重新优化抗原去隐蔽方案。
所有载玻片上染色杂乱显色物质衰退	• 改变脱蜡试剂。 • 再脱水的最后步骤使用二甲苯。 • 尝试另外的二甲苯替代物。 • 检查水性封固剂（如果使用的话）与显色物质的相容性。

IHC 和通常的组织学讨论组可以在因特网上找到。histonet 资料库（http://www.histosearch.com/histonet），IHC 主页（http://immuno.hypermart.net），组织学讨论邮件清单页面（http://www.histology.to/listservers.html），和 histonet 讨论邮件清单（http://www.histosearch.com/listserver.html）提供了关于疑难解答有价值的信息，

以及购买的窍门和推荐的试剂。抗体生产商制作了 IHC 操作的小册子，其内容包含了详细的疑难解答部分（Naish 1989；Chemicon International 1998）。

643

替代方案：抗原去隐蔽和胰蛋白酶消化

下面提供了两种方法用于抗原去隐蔽：在压力锅中加热载玻片和用胰蛋白酶处理它们。

用于抗原恢复的压力锅方法

本方法通过电热台加热不锈钢压力锅，载玻片在压力锅中的柠檬酸盐缓冲液（10 mmol/L，pH 6.0）里加热。缓冲液的用量随压力锅的尺寸而变化。确保有足够的缓冲液覆盖全部组织切片。每次抗原恢复最多处理 48 片载玻片，以帮助各次免疫组织化学实验达成一致。在我们实验室中，柠檬酸盐缓冲液（pH 6.0）对一系列抗体是抗原恢复理想的缓冲液。

该压力锅方法可以用于其他抗原恢复缓冲液，如 EDTA 和 Tris-HCl，作为柠檬酸盐缓冲液的替代以确定所使用的每种抗体最佳抗原恢复方法（表 8-4）（Shi et al. 1995；Pileri et al. 1997；McNichol and Richmond 1998）。

附加材料

柠檬酸钠（10 mmol/L，pH 6.0）
压力锅

方法

1. 将 2 L 10 mmol/L 柠檬酸钠（pH 6.0）放入上面有盖的压力锅中。
   根据所使用压力锅的尺寸，10 mmol/L 柠檬酸钠的用量需作相应调整。
   确保抗原恢复缓冲液覆盖全部组织阵列的切片。
2. 用电热板的最高设定加热缓冲液直至沸腾。
3. 在压力锅中将载玻片（已脱蜡和重新水化）浸入沸腾的缓冲液。
4. 更换锅盖并关上压力锅。
5. 根据制造商的说明将压力锅升压至合适的压力。
6. 当压力锅达到完全压力时开始计时 2 min。
   对于一些抗体或者组织阵列样品，最佳染色可能需要延长抗原恢复的时间。这要根据经验来检查。
7. 在 2 min 的孵育时间结束时，根据制造商的说明，小心的降低压力锅的压力。
8. 打开锅盖，将载玻片放入装有自来水的器皿中。
9. 将载玻片冷却 20 min，然后继续免疫组织化学染色。

胰蛋白酶消化

含 0.1%胰蛋白酶的 0.1%氯化钙溶液用于消化组织切片中，在固定时形成的蛋白交联。酶消化在 37 ℃下进行。

方法

1. 将一个装有 50 ml 蒸馏水的 Coplin 广口瓶，和一个装有 50 ml 0.1%（*m*/*V*）氯化钙溶液的 Coplin 广口瓶放入 37 ℃的水浴中。
2. 使用温度计检查两种溶液的温度为 37 ℃。
3. 将载玻片在加热的蒸馏水中放置 10 min。
4. 加入 0.5 ml 解冻的 10%胰蛋白酶到装有 0.1%（*m*/*V*）氯化钙溶液的 Coplin 广口瓶中。
5. 使用 50 mmol/L 氢氧化钠调节加热的氯化钙/胰蛋白酶至 pH 7.8（用 pH 计测试）。
6. 将氯化钙/胰蛋白酶回到 37 ℃。
7. 将载玻片放入氯化钙/胰蛋白酶溶液中 10 min。
8. 在完成消化后将载玻片在流动的自来水中清洗 3 min。
9. 继续免疫组织化学操作。

参考文献

Angerer L.M., Cox K.H., and Angerer R.C. 1987. Demonstration of tissue-specific gene expression by in situ hybridization. *Methods Enzymol.* **152:** 649–661.

Bales K.R., Hannon K., Smith C.K., 2nd, and Santerre R.F. 1993. Single-stranded RNA probes generated from PCR-derived DNA templates. *Mol. Cell Probes* **7:** 269–275.

Bancroft J.D. and Stevens A.S. 1996. *Theory and practice of histological techniques*, Churchill Livingstone, London.

Barlund M., Monni O., Kononen J., Cornelison R., Torhorst J., Sauter G., Kallioniemi O.-P., and Kallioniemi A. 2000. Multiple genes at 17q23 undergo amplification and overexpression in breast cancer. *Cancer Res.* **60:** 5340–5344.

Bobrow M.N., Litt G.J., Shaughnessy K.J., Mayer P.C., and Conlon J. 1992. The use of catalyzed reporter deposition as a means of signal amplification in a variety of formats. *J. Immunol. Methods* **150:** 145–149.

Bubendorf L., Kononen J., Koivisto P., Schraml P., Moch H., Gasser T.C., Willi N., Mihatsch M.J., Sauter G., and Kallioniemi O.P. 1999. Survey of gene amplifications during prostate cancer progression by high-throughput fluorescence in situ hybridization on tissue microarrays. *Cancer Res.* **59:** 803–806.

Camp R.L., Charette L.A., and Rimm D.L. 2000. Validation of tissue microarray technology in breast carcinoma. *Lab. Invest.* **80:** 1943–1949.

Chemicon International Inc. 1998. *Introduction to antibodies*. Chemicon International, Temecula, California.

Cox K.H., DeLeon D.V., Angerer L.M., and Angerer R.C. 1984. Detection of mRNAs in sea urchin embryos by in situ hybridization using asymmetric RNA probes. *Dev. Biol.* **101:** 485–502.

Frost A.R., Sparks D., and Grizzle W.E. 2000. Methods of antigen recovery vary in their usefulness in unmasking specific antigens in immunohistochemistry. *Appl. Immunohistochem. Mol. Morphol.* **8:** 236–243.

Gillett C.E., Springall R.J., Barnes D.M., and Hanby A.M. 2000. Multiple tissue core arrays in histopathology research: a validation study. *J. Pathol.* **192:** 549–553.

Hedenfalk I., Duggan D., Chen Y., Radmacher M., Bittner M., Simon R., Meltzer P., Gusterson B., Esteller M., Kallioniemi O.P., Wilfond B., Borg A., and Trent J. 2001. Gene-expression profiles in hereditary breast cancer. *N. Engl. J. Med.* **344:** 539–548.

Heiles H.B., Genersch E., Kessler C., Neumann R., and Eggers H.J. 1988. In situ hybridization with digoxigenin-labeled DNA of human papillomaviruses (HPV 16/18) in HeLa and SiHa cells. *BioTechniques* **6:** 978–981.

Holtke H.J. and Kessler C. 1990. Non-radioactive labeling of RNA transcripts in vitro with the hapten digoxigenin (DIG); hybridization and ELISA-based detection. *Nucleic Acids Res.* **18:** 5843–5851.

645 Hoos A., Urist M.J., Stojadinovic A., Mastorides S., Dudas M.E., Leung D.H., Kuo D., Brennan M.F., Lewis J.J., and Cordon-Cardo C. 2001. Validation of tissue microarrays for immunohistochemical profiling of cancer specimens using the example of human fibroblastic tumors. *Am. J. Pathol.* **158:** 1245–1251.

Hopwood D. 1996. Fixation and fixatives. In *Theory and practice of histological techniques* (ed. J.D. Bancroft and A.S. Stevens), pp. 23–46. Churchill Livingstone, London.

John H.A., Birnstiel M.L., and Jones K.W. 1969. RNA-DNA hybrids at the cytological level. *Nature* **223:** 582–587.

Kallioniemi O.-P., Wagner U., Konenen J., and Sauter G. 2001. Tissue microarray technology for high-throughput molecular profiling of cancer. *Hum. Mol. Genet.* **10:** 657–662.

Kessler C., Holtke H.J., Seibl R., Burg J., and Muhlegger K. 1990. Non-radioactive labeling and detection of nucleic acids. I. A novel DNA labeling and detection system based on digoxigenin: Anti-digoxigenin ELISA principle (digoxigenin system). *Biol. Chem. Hoppe Seyler* **371:** 917–927.

King G., Payne S., Walker F., and Murray G.I. 1997. A highly sensitive detection method for immunohistochemistry using biotinylated tyramine. *J. Pathol.* **183:** 237–241.

Komminoth P. 1992, Digoxigenin as an alternative probe labeling for in situ hybridization. *Diagn. Mol. Pathol.* **1:** 142–150.

Komminoth P., Merk F.B., Leav I., Wolfe H.J., and Roth J. 1992. Comparison of 35S- and digoxigenin-labeled RNA and oligonucleotide probes for in situ hybridization. Expression of mRNA of the seminal vesicle secretion protein II and androgen receptor genes in the rat prostate. *Histochemistry* **98:** 217–228.

Kononen J., Bubendorf L., Kallioniemi A., Barlund M., Schraml P., Leighton S., Torhorst J., Mihatsch M.J., Sauter G., and Kallioniemi O.P. 1998. Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat. Med.* **4:** 844–847.

Lipshutz R.J., Fodor S.P., Gingeras T.R., and Lockhart D.J. 1999. High density synthetic oligonucleotide arrays. *Nat. Genet.* (suppl.) **21:** 20–24.

McNicol A.M. and Richmond J.A. 1998. Optimizing immunohistochemistry: Antigen retrieval and signal amplification. *Histopathology* **32:** 97–103.

Miller K. 1996. Immunocytochemical techniques. In *Theory and practice of histological techniques* (ed. J.D. Bancroft, and A. S. Stevens, Churchill Livingstone, London. 435–470.

Miller R.T., Swanson P.E., and Wick M.R. 2000. Fixation and epitope retrieval in diagnostic immunohistochemistry: A concise review with practical considerations. *Appl. Immunohistochem. Mol. Morphol.* **8:** 228–235.

Moch H., Kononen T., Kallioniemi O.P., and Sauter G. 2001. Tissue microarrays: What will they bring to molecular and anatomic pathology? *Adv. Anat. Pathol.* **8:** 14–20.

Muhlegger K., Huber E., von der Eltz H., Ruger R., and Kessler C. 1990. Non-radioactive labeling and detection of nucleic acids. IV. Synthesis and properties of digoxigenin-modified 2′-deoxyuridine-5′-triphosphates and a photoactivatable analog of digoxigenin (photodigoxigenin). *Biol. Chem. Hoppe Seyler* **371:** 953–965.

Naish S.J. 1989. *Handbook: Immunochemical staining methods.* Dako Corporation, Carpinteria.

Nocito A., Bubendorf L., Tinner E.A., Suess K., Wagner U., Forster T., Kononen J., Fijan A., Bruderer J., Schmid U., Ackermann D., Maurer R., Alund G., Knonagel H., Rist M., Anabitarte M., Hering F., Hardmeier T., Schoenenberger A.J., Flury R., Jager P., Luc Fehr J., Schraml P., Moch H., Mihatsch M.J., Gasser T., and Sauter G. 2001. Microarrays of bladder cancer tissue are highly representative of proliferation index and histological grade. *J. Pathol.* **194:** 349–357.

Pardue M.L. and Gall J.G. 1969. Molecular hybridization of radioactive DNA to the DNA of cytological preparations. *Proc. Natl. Acad. Sci.* **64:** 600–604.

Pileri S.A., Roncador G., Ceccarelli C., Piccioli M., Briskomatis A., Sabattini E., Ascani S., Santini D., Piccaluga P.P., Leone O., Damiani S., Ercolessi C., Sandri F., Pieri F., Leoncini L., and Falini B. 1997. Antigen retrieval techniques in immunohistochemistry: Comparison of different methods. *J. Pathol.* **183:** 116–123.

Rimm D.L., Camp R.L., Charette L.A., Costa J., Olsen D.A., and Reiss M. 2001. Tissue microarray: A new technology for amplification of tissue resources. *Cancer J.* **7:** 24–31.

Sallinen S.L., Sallinen P.K., Haapasalo H.K., Helin H.J., Helen P.T., Schraml P., Kallioniemi O.P., and Kononen J. 2000. Identification of differentially expressed genes in human gliomas by DNA microarray and tissue chip techniques. *Cancer Res.* **60:** 6617–6622.

Sambrook J. and Russell D. 2001. *Molecular cloning: A laboratory manual,* 3rd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Shi S.-R., Cote R.J., Chaiwun B., Young L.L., Shi Y., Hawes D., Chen T., and Taylor C.R. 1998. Standardization of immunohistochemistry based on antigen retrieval technique for antigen retrieval technique for routine formalin-fixed tissue sections. *Appl. Immunohistochem.* **6:** 89–96.

Shi S.R., Cote R.J., and Taylor C.R. 1997. Antigen retrieval immunohistochemistry: Past, present, and future. *J. Histochem. Cytochem.* **45:** 327–343.

Shi S.R., Imam S.A., Young L., Cote R.J., and Taylor C.R. 1995. Antigen retrieval immunohistochemistry under the influence of pH using monoclonal antibodies. *J. Histochem. Cytochem.* **43:** 193–201.

Taylor C.R., Shi S.-R., and Cote R.J. 1996. Antigen retrieval for immunohistochemistry. *Appl. Immunohistochem.* **4:** 144–166.

Wester K., Wahlund E., Sundstrom C., Ranefall P., Bengtsson E., Russell P.J., Ow K.T., Malmstrom P.U., and Busch C. 2000. Paraffin section storage and immunohistochemistry. Effects of time, temperature, fixation, and retrieval protocol with emphasis on p53 protein and MIB1 antigen. *Appl. Immunohistochem. Mol. Morphol.* **8:** 61–70.

Wetmur J.G. 1991. DNA probes: Applications of the principles of nucleic acid hybridization. *Crit. Rev. Biochem. Mol. Biol.* **26:** 227–259.

Wetmur J.G. and Davidson N. 1968. Kinetics of renaturation of DNA. *J. Mol. Biol.* **31:** 349–370.

Wilkinson D.G. 1992. Theory and practice of in situ hybridization. In *In situ hybridization: A practical approach* (ed. D.G. Wilkinson). IRL Press, Oxford, United Kingdom.

Yang H., Wanner I.B., Roper S.D., and Chaudhari N. 1999. An optimized method for in situ hybridization with signal amplification that allows the detection of rare mRNAs. *J. Histochem. Cytochem.* **47:** 431–446.

因特网资源

http://www.histosearch.com/histonet The Histonet Archives.

http://immuno.hypermart.net The Immunohistochemistry homepage: A comprehensive list of immunohistochemistry resources on the Internet.

http://www.histology.to/listservers.html Histology Listerservers and Newsgroups.

http://www.histosearch.com/listserver.html The Histonet Listserver.

附录

附录 1* 构建全长基因文库

Piero Carninci
(基因组科学实验室，Riken Main Campus，和光市，埼玉县，日本)

虽然富含全长 cDNA 的代表性文库的制备很困难，但是在使用中的效果却令人满意。这类文库对于应用 EST 单向测序技术发现大规模基因的项目而言是特别有价值的。当文库内相关的 DNA 克隆中同时包含有完整的编码序列和未转录的 3′与 5′区域时，下游分析包括蛋白质表达、生物信息学和功能指定将会被极大地简化。但是，在传统上有一些问题阻碍了全长 cDNA 的产生，以及生成以相等的频率表示全部 cDNA 的文库。这些问题有：

- 反转录酶对哺乳类动物 cDNA 的 5′区域的拷贝效率较低。稳定的二级结构使得拷贝过程过早终止，并且酶分子从很多的 mRNA 模板上发生脱离，导致了在 cDNA 克隆中缺少对应于 mRNA 5′端区域的序列。当使用寡聚 dT 作为引物进行第一链合成时，这一问题会更加严重。
- 缺少从第一链 cDNA 合成产物中选择性地去除被截短的 cDNA（truncated DNA）的有效方法。
- 没有研究出能够以相同的效率来插入和扩增多种长度 cDNA 的载体。当前使用的大多数载体对较短的插入片段的扩增更为有效，由此导致了对较长 cDNA 的选择性淘汰。
- 构建各种 mRNA 均以相同频率出现的均一化 cDNA 文库时遇到的困难。生成代表性文库时要求对最初的 mRNA 分子群中各种 mRNA 分子之间丰度的巨大差异进行校正，并且在文库的克隆和扩增过程中不会引入额外的偏差，例如，在通过 PCR 扩增 cDNA 序列时或者是在生长基于质粒的文库时。

在过去的数年中，我们研究了一个可以克服一些上述问题的方案：

1. 使用 Cap-trapping 方法来选择全长的 cDNA（Carninci and Hayashizaki 1999）。将生物素残基通过化学的方法连接到位于哺乳动物 mRNA 5′端的帽子结构上的二醇基团。RNase Ⅰ用于从未完整合成的 cDNA-mRNA 杂交体中消除生物素化的帽子结构。全长的 cDNA-mRNA 杂交体可以通过与包被了链亲和素的磁珠结合而被回收。

---

* 本附录是由 Siân Curtis 根据 Piero Carninci 供稿的材料编写而成，其中那些标准的或者是近似标准的方法，请读者参考公开发表的资料。凡有重要修改/改进的地方，在表格后面概述了可供选择的方案。

2. 在第一链 cDNA 合成时，用海藻糖来稳定反转录酶（Carninci et al. 1998；Carninci and Hayashizaki 1999）。
3. 在构建代表性全长 cDNA 文库时，选用能够有效地插入和扩增长链 cDNA 的克隆载体非常重要。否则，在文库中对这类 cDNA 的代表性会严重地降低，甚至在最坏的情形下可能从文库中缺失。为了消除基于分子长短的差异，要选择可以容纳和有效扩增长链 cDNA（>15 kb）的大容量噬菌体载体来克隆 cDNA。除了具有大的容量外，载体必须具备可以从 cDNA 文库中以质粒的形式完整地切除（bulk-excision）插入片段的序列。综合了这种扩增和切除全长 cDNA 文库性能的最佳载体是 λ-FLC 家族，它们包含：①容纳广泛长度的真核 DNA 的"分子长度平衡机制（size balancer）"；②用于直接进行质粒完整切除的两套切除系统（Cre-*lox*P 和 Gateway）；③可以将 cDNA 转导到具有特殊功能载体中的限制性酶切位点（Carninci et al. 2001）。
4. 新基因发现过程中的许多问题源自哺乳动物细胞中 mRNA 之间表达水平的差异。哺乳动物的总 mRNA 包含了三种或者更多的动力学类型，它们的丰度各不相同。最占优势的转录本占到 mRNA 质量的 30%～90%，但是只占到 mRNA 序列复杂度的 0.1%；中等丰度的 mRNA 约占 mRNA 质量的 50%～60%和序列复杂度的 2%～5%。稀有 mRNA 在质量中所占的比例微乎其微，但是却提供了大于 95%的多核糖体 RNA 的序列复杂度。因为这些丰度上的差异，从常规的 cDNA 文库中随机选择 cDNA 进行单向测序是一种效率非常低的分离低表达基因的方法。

通过降低超高丰度和中等程度表达的 cDNA 的优势，并使得所有种类的 cDNA 以相似的频率出现，可以提高单向测序的效率。从最终的文库中扣除不需要的 cDNA，可以使得效率得到进一步的提高。在本方案中，通过仔细控制的第一链 cDNA 与 RNA 驱动子（RNA driver）的杂交来实现均一化和扣除过程。在进行均一化时，驱动子是最初用作 cDNA 合成模板的那部分生物素化的 RNA；在进行扣除时，驱动子是从已鉴定的 cDNA 克隆集通过体外转录制备的生物素化的 RNA 分子群。高丰度的和不需要的 cDNA 与驱动子之间形成的杂交体被吸附到链亲和素包被的磁珠上而除去，得到单链的、经过标准化和扣除的 cDNA 分子群。

本方案的组织

本方案分成 17 个阶段（Ⅰ～ⅩⅦ），从模板 mRNA 的抽提和纯化开始，结束于对 cDNA 文库的评价（图 A1-1）。接下来的内容对所有的阶段和方法进行了总结，以便于读者确定生成全长 cDNA 文库所包含的技术范围和需要的工作量。不过，这个总结并不意味着它可以被当作实验室中的详细指南。对构建全长 cDNA 文库感兴趣的读者应该按照本书的网站 www.dna-microarrays.com 上完整的介绍来进行实验。

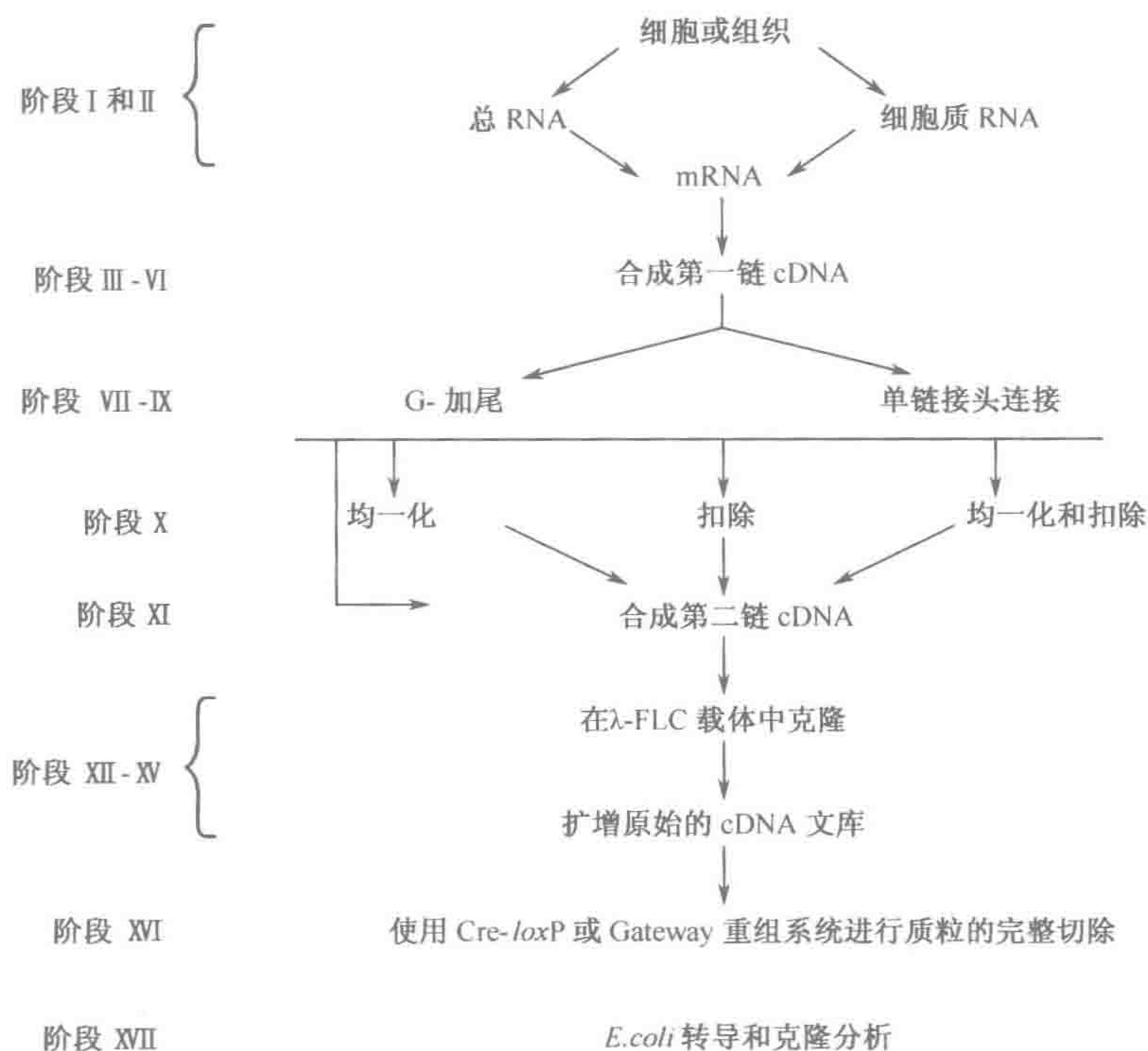


图 A1-1　构建均一化的全长 cDNA 文库中各阶段的流程图

649

表 A1-1　全长 cDNA 文库的构建

阶　段	方　法	修改/注释
阶段Ⅰ：RNA 的制备	抽提和纯化 用异硫氰酸胍抽提	见 Sambrook 和 Russell（2001），Carninci 和 Hayashizaki（1999），Carninci 等（2002） 抽提后，使用 CTAB 沉淀法去除多糖［包括 Sambrook 和 Russell（2001）第 7 章的方案 1 中步骤 10 以后的操作］。要注意 hnRNA 出现在最终的产物中（含量约达到 5%）［Sambrook 和 Russell（2001），第 7 章，方案 1］（见本表后的附加方案）
	用 NP 40 裂解和 CTAB-尿素沉淀	使用 NP40 在含有氧钒核糖核苷复合物的缓冲液中裂解细胞，回收细胞质中的 RNA。在最终产物中不存在 hnRNA（方法见附加方案）
阶段Ⅱ：mRNA 的纯化	选取 poly(A)$^+$ mRNA	应用基于寡聚(dT)-纤维素色谱法的商品化试剂盒纯化 mRNA（Sambrook and Russell 2001，第 7 章，方案 3；同时参考供应商的手册）
阶段Ⅲ：引物的设计和制备	设计时的考虑，以及用来纯化合成的寡核苷酸的方法	建议引物上带有甲基化敏感的限制性内切核酸酶的识别位点，以便于后面进行克隆时在全长分子的末端进行选择性的断裂。寡核苷酸引物按照标准的方法进行合成和纯化。引物的最终纯化通过变性聚丙烯酰胺凝胶电泳来完成（Sambrook and Russell 2001，第 10 章，导言和方案 1；同时参考第 10.48 页“纯化合成 Oligos 的方法”）

续表

阶　段	方　法	修改/注释
阶段Ⅳ：第一链的合成	在海藻糖和山梨醇存在下进行全长 cDNA 的合成	使用 SUPERSCRIPT Ⅱ反转录酶（Life Technologies/ Invitrogen）进行第一链的合成，基本上如 Sambrook 和 Russell（2001，第 11 章，方案 1）所述 对合成第一链反应所作的修改包括： 1. 用 5′-甲基 dCTP 代替 dCTP 掺入，生成甲基化的 cDNA。由于在反应中使用了 5′-甲基 dCTP，在平行的小规模反应中使用［α-^{32}P］dGTP 而非［α-^{32}P］dCTP 作为放射性标记。 2. 反应中加入海藻糖和山梨醇，用于在高反应温度（约 56℃）下稳定反转录酶（Carninci et al. 1998，2000）。 3. "热启动"反应可以使 mRNA 中棘手的二级结构容易地转录成全长 cDNA。
	在带有海藻糖和山梨醇的 GCI 缓冲液中合成全长 cDNA	此处第一链合成方案是对阶段Ⅳ的修改。反应中使用了 GCI 缓冲液（TaKaRa），它是为在长链 PCR 中替代标准的 5×第一链缓冲液而优化的
阶段Ⅴ：测量样品产量	测量核酸的放射性	第一链反应的产量使用 DE-81 吸收进行计算（Sambrook and Russell，2001，附录 8，第 A8.26 页） 第一链 cDNA 产物的大小在琼脂糖凝胶上进行观察（Sambrook and Russell 2001，第 5 章，方案 8）
阶段Ⅵ：沉淀第一链 cDNA	产物的 CTAB 沉淀	RNA 的 CTAB 沉淀去除了影响 Cap-trapping 过程的蛋白质和多糖（方法见附加方案）
阶段Ⅶ：富集全长 cDNA	5′帽子结构的氧化和生物素化	通过对 mRNA 5′端帽子结构的生物素化修饰使得第一链 cDNA 产物中的全长 cDNA 得到富集（方法见附加方案；Carninci and Hayashizaki 1999）
阶段Ⅷ：富集全长 cDNA	全长 cDNA 的 Cap-trapping 和释放	通过对 5′端帽子结构的生物素化处理，可以使用链亲和素包被的磁珠对全长 cDNA 进行选择（方法见附加方案；Carninci and Hayashizaki 1999）
阶段Ⅸ：第二链的合成	向 cDNA 的 5′端添加引导位点	有多种方法可以用来引导第二链的合成（见 Sambrook and Russell 2001，第 11 章导言）。建议两种替代方法（方法 1 更为常用）： 1. 单链接头连接方法（SSLLM）向单链 cDNA 添加一个双链的 DNA 接头。该接头带有随机 6 bp（dN6 或 dGN5）的 3′突出端，能与任何 cDNA 序列发生连接。本方法简化了从 cDNA 克隆进行 cDNA 测序和蛋白质表达（SSLLM 方法见附加方案）。 2. cDNA 的 3′端进行 dG 残基的同聚物加尾（Deininger 1987）。然后使用寡聚（dC）合成引物引导第二链的合成（Sambrook and Russell 2001，第 11.17～11.18）。

阶 段	方 法	修改/注释
阶段Ⅹ：cDNA 文库的均一化和扣除		当复杂或稀少的 cDNA 为单向测序所鉴别时，高丰度的 cDNA 极大地削弱了发现所感兴趣克隆的概率。为了增加单向测序的效率，使用基于核酸杂交动力学的技术可以减小高水平表达 cDNA 的代表性。"均一化"可使得高丰度和低丰度的种类在 cDNA 文库中得到同等的表示。均一化和扣除的详细信息，请参见 Carninci 等（2000）以及 Hames 和 Higgins（1985）。除了对文库进行均一化之外，消除不需要的 cDNA 也是非常有用的。将从感兴趣组织中所抽提 mRNA 制备的单链 cDNA，与过量的、从不表达所感兴趣基因的其他来源制备的 DNA 或 RNA（驱动子）杂交可以实现这个目的。先前已经过鉴定的克隆也可以用作消减驱动子，以减少对同一克隆重复识别的概率。（Sambrook and Russell 2001，第 11 章导言）
阶段Ⅺ：第二链 cDNA 合成		经过 SSLLM 处理，和（或）均一化与扣除的第一链 cDNA 用作第二链 cDNA 合成的底物。使用长链热稳定聚合酶的混合物，通过引物延伸进行第二链的合成，这样便于有效地合成长链的第二链 cDNA［长链 PCR 方法见附加方案（Cheng et al. 1994）；同时见 Sambrook and Russell 2001，第 8 章，方案 13］
阶段Ⅻ：cDNA 的断裂	制备用于克隆的 cDNA	由于甲基-dCTP 掺入了第一链 cDNA，所以双链 cDNA 是半甲基化的。但是 cDNA 两侧的接头引物中却不是这样。这就提供了一种完全消化两侧序列而不会断裂 cDNA 序列的简单方法。本方案中使用包含了 *Xho* Ⅰ 和 *Bam*H Ⅰ 位点的引物。步骤如下： 1. 限制性消化：为保证消化完全，进行两个独立的反应，每个反应中使用一种酶（在本例中，*Xho* Ⅰ 和 *Bam*H Ⅰ），要使用过量的酶（25～50 单位/μg cDNA）。每次反应后用蛋白酶 K 消化，酚-氯仿抽提〈!〉和乙醇沉淀来处理样品（Sambrook and Russell 2001，附录 8）。 2. 去磷酸化：为避免克隆多个插入片段，建议对 cDNA 作去磷酸化处理（Sambrook and Russell 2001，第 9 章，方案 13）。如果载体已经是去磷酸化的，cDNA 不能作去磷酸化处理。
阶段ⅩⅢ：按大小分离双链 cDNA	用 CL-4B 或琼脂糖凝胶电泳分离	为了除去所有的引物接头或者短的 cDNA，克隆之前要按分子大小对 cDNA 进行分离。建议使用 CL-4B 离心柱或琼脂糖凝胶（Sambrook and Russell 2001，第 11 章，方案 1，阶段 5，或者第 5 章，方案 1）
阶段ⅩⅣ和ⅩⅤ：克隆载体的准备和使用	阶段ⅩⅢ：琼脂糖凝胶分离（替代 CL-4B）	许多载体都适于 cDNA 的克隆。建议使用 λ-FLC 家族载体（Carninci et al. 2001），因为它可以容纳大于 15 kb 的 cDNA，并且对于小片段没有偏性。关于对 λ-FLC 的介绍，见 Carninci et al.（2001） 载体的准备、检测和连接方法，见 Sambrook 和 Russell（2001，第 11 章，方案 1，阶段 6）

续表

阶 段	方 法	修改/注释
阶段 XVI：插入片段的完整切除	在体的，固相剪切和 λ-FLC 介导的完整切除	克隆载体的切除机制，尤其是 λZap 和 λZipLox 系列，见 Sambrook 和 Russell（2001，第 11 章导言，第 11.34）的讨论。关于 λ-FLC 载体的使用（Carninci et al. 2001），在网站上的补充材料中有详细的说明
阶段 XVII：对 cDNA 文库的评价		鉴定 cDNA 克隆的方法见 Sambrook 和 Russell（2001，第 11 章导言，第 11.34） 对 cDNA 文库中插入片段的大小进行评价时，选取 20～30 个克隆，用裂解液进行长链 PCR，或者为每个克隆进行小规模的 DNA 制备，并用限制性内切核酸酶消化。使用凝胶电泳对结果进行分析。全长 cDNA 的比例可以通过末端测序，并与公共数据库中的已知基因进行比较而得出

附加方案

652

阶段 I：RNA 的制备

使用异硫氰酸胍从细胞中抽提 RNA

应用于 DNA 阵列时，建议使用 CTAB 沉淀步骤以除去多糖（包括在第 7 章，方案 1，步骤 10 以后，Sambrook and Russell 2001）。本方法操作迅速；但是最终产物中存在核不均一 RNA（hnRNA）（达到约 5%）。关于 RNA 的质量控制，见 Sambrook 和 Russell（2001，第 7 章中信息栏）。

方法

1. 按照 Sambrook 和 Russell（2001）第 7 章，方案 1 操作，至步骤 10。
2. 使用 2 倍体积的 CTAB/尿素溶液（1% CTAB，4 mol/L 尿素，50 mmol/L Tris-HCl，pH 7.0，1 mmol/L EDTA，pH 8.0）沉淀样品。
3. 在 4 ℃下，以 9500 *g* 离心 15 min，回收 RNA。
4. 将沉淀溶解于 4 ml 7 mol/L 的盐酸胍中。
5. 用异丙醇沉淀样品，并将团块在 500～1000 μl H_2O 中重悬。

用 NP40 缓冲液从细胞中抽提 RNA

作为替代，也可以用含 NP40 的缓冲液处理细胞，然后用 CTAB/尿素沉淀。在这种方法中，最终产物中不带有 hnRNA。关于 RNA 的质量控制，见 Sambrook 和 Russell（2001，第 7 章中信息栏）。

方法

注意：标记有〈!〉的材料的正确处理方法请参见附录 3。

1. 收集细胞，并重悬于 2 ml 不含 Ca^{2+} 和 Mg^{2+} 的磷酸盐缓冲液（PBS）中。
2. 加入 5 ml 含有 20 mmol/L 氧钒核糖核苷复合物的裂解缓冲液（100 mmol/L NaCl，5 mmol/L $MgCl_2$，50 mmol/L Tris-HCl，pH 7.5，0.5% NP40），裂解细胞。
3. 充分混匀溶液，在冰上孵育 3 min。
4. 在 4 ℃下，将溶液以 9500 *g* 离心 1～3 min。包含 DNA 和 hnRNA 的核将从溶液中以团块方式析出，残存的细胞质 RNA 留在上清液中。
5. 将上清液转移到含有 2 倍体积 CTAB/尿素溶液（1% CTAB，4mol/L 尿素，50 mmol/L Tris-HCl，pH 7.0，1 mmol/L EDTA，pH 8.0）的试管中。
6. 在 4 ℃下，以 9500 *g* 离心 10 min，回收 RNA。
7. 将团块溶解于 4 ml 7mol/L 的盐酸胍中。
8. 用酚-氯仿〈!〉抽提和异丙醇沉淀纯化样品，并将团块重悬在 100～500 μl H_2O 中。

阶段Ⅵ：第一链 cDNA 的 CTAB 沉淀

以蛋白酶 K 消化增效 CTAB 沉淀是用来纯化样品、除去蛋白质和多聚糖污染物的，它们可干扰步骤Ⅶ和Ⅷ的帽子结构结合方法富集过程。

方法

注意：标记有〈!〉的材料的正确处理方法请参见附录 3。

警告：本阶段将处理具有高放射性的样品。在未掺入的 dNTP 去除之前，应予以高度重视。

1. 将大规模的 cDNA 第一链反应和包含 [α-^{32}P] dGTP〈!〉的小规模试验性反应混合。
2. 加入 3 μl EDTA（0.5 mol/L，最终浓度为 10 mmol/L）和 2 μl 蛋白酶 K（10 μg/μl）。将体积调整到 150 μl，在 45～50 ℃下孵育 15 min 至 1 h。
3. 孵育结束后，加入 32 μl 氯化钠（5 mol/L，不含 RNase），320 μl CTAB/尿素溶液 [1% CTAB，4 mol/L 尿素，50 mmol/L Tris-HCl，pH 7.0，1 μmol/L EDTA（0.5 mol/L 时 pH 8.0）]。在室温下再孵育 10min。
4. 以 15 000 *g*（或是微量离心机的最高转速）离心 10 min，小心弃去上清液，其中包含了大部分游离的 dNTP，具有放射性〈!〉。
5. 小心地将团块溶解在 100 μl 盐酸胍（7 mol/L）中，加入 250 μl 乙醇，在冰上（或－20 ℃、－80 ℃）放置 30～60 min（少量 cDNA 需要的时间更长）。
6. 以 15 000 *g*（或是微量离心机的最高转速）离心 10 min，小心弃去上清液。使用 800 μl 80%乙醇非常小心地洗涤团块 2 次。再次在 15 000 *g* 下离心 3 min。
7. 将 cDNA 溶解在 46 μl H_2O（非 TE）中，继续阶段Ⅶ和Ⅷ中的富集方案。

   此处给出的体积与第一链 cDNA 反应体积成比例。当起始体积不同时，应按比例进行适当的缩放。

   使用手持计数器检查试管和溶液的计数，确保团块已经完全重悬。完全重悬时，管壁上残

653

留的放射性应低于总记数的5%。

如果cDNA的合成规模非常小（<2 μg起始mRNA），可以加入载体以提高沉淀效率。不能使用糖原。

阶段Ⅶ：富集全长cDNA：5′帽子结构的氧化和生物素化

654

第一链cDNA合成后，通过对mRNA 5′端帽子结构的氧化和生物素化来富集产物中全长的cDNA。对5′端帽子结构生物素化后，就可以按照阶段Ⅷ中介绍的方法（Carninci and Hayashizaki 1999）应用链亲和素标记的磁珠来分离全长cDNA。在本方案中，mRNA的二醇基团在碘酸钠存在下被氧化，随后被生物素化。

方法

注意： 标记有〈!〉的材料的正确处理方法参见附录3。

警告： 所有溶液必须是不含RNase的。使用检测过的试剂和无菌H_2O，在不含RNase的玻璃或者塑料容器中配制溶液。乙酸盐和柠檬酸盐缓冲液可以用高压灭菌。

mRNA二醇基团的氧化

1. 将下列试剂直接加入阶段Ⅵ中得到的46 μl CTAB沉淀的第一链反应中。

   3.3 μl 1 mol/L 醋酸钠（pH 4.5）

   100 mmol/L $NaIO_4$（新鲜配制的）至最终浓度为10 mmol/L

   放置在暗处，在冰上孵育45min。加入1 μl 80%的甘油并涡旋振荡以终止反应。

2. 用61 μl异丙醇、0.5 μl 10% SDS〈!〉和11 μl 5 mol/L NaCl沉淀，回收核酸。在暗处，于－20 ℃或－80℃下孵育30 min，然后以15 000 g离心15 min。小心弃去上清液。

   从这里开始，大多数放射性计数应该存在于乙醇沉淀物中。如果计数仍然在上清液中，沉淀就不完全。

3. 用500 μl 80%乙醇清洗团块，然后以15 000 g离心2～3 min。小心弃去上清液。

4. 再次短暂离心，除去任何残留的乙醇。将核酸溶解在50μl H_2O中。

氧化后二醇基团的衍生

5. 将0.56 g生物素酰肼溶解在160 μl新鲜配制的生物素化缓冲液［33 mmol/L柠檬酸钠，pH 6.0（来自1 mol/L储存液）］，0.33% SDS［来自10%（*m/V*）储存液］中，配制浓度为10 mmol/L生物素酰肼（修饰了长手臂分子的类型）溶液。

6. 向步骤4得到的第一链cDNA（50 μl）中加入160 μl缓冲液，轻轻混匀。在室温下孵育过夜（10～16 h），或者在37 ℃下孵育3～4 h。

   有些批次的生物素酰肼在37 ℃下会导致DNA发生降解。

655 7. 孵育结束时，加入以下试剂沉淀生物素化的 cDNA-mRNA 杂交体：

1 mol/L 柠檬酸钠（pH 6.1）	75 μl
5 mol/L NaCl	5 μl
乙醇	750 μl

轻轻混匀，在冰上孵育 1 h，或者在−80 ℃或−20 ℃下孵育 30 min。

8. 以 15 000 *g* 的速度离心样品 10 min。弃去上清液，用 80%乙醇清洗 3 次，每次清洗后离心处理，以从团块中去除游离的生物素酰肼。
9. 将最终得到的团块溶解在 75 μl TE 中。使用手持式计数器和肉眼观察，用于检查重悬是否完全。

阶段Ⅷ：富集全长 cDNA：CAP-TRAPPING 和全长 cDNA 的释放

在进行了阶段Ⅶ中叙述的生物素化反应后，用 RNase Ⅰ处理样品，它能够特异性地降解单链 RNA。本反应作用于不完整的异源双链分子的 5′帽子结构，并除去 3′多聚(A) 加尾的单链部分。然后，生物素化的分子被链亲和素包被的磁珠所俘获，并在适合于保留小分子和释放长链的 cDNA 的条件下进行洗脱（Kotsopoulos and Shuber 1996)。本方案中推荐使用链亲和素包被的磁性多孔玻璃（MPG）珠（CPG Inc.，New Jersey)。为了将核酸的非特异性结合，尤其是非全长 cDNA 的结合程度减小到最低，磁珠事先要在过量的、不含 DNA 的 tRNA 中预孵育处理。

方法

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

1. 对上面步骤 9 得到的 cDNA 样品中（75 μl）加入：

RNase Ⅰ缓冲液（Promega）	20 μl
RNaseⅠ（5 或 10 μl/μl，RNase Ⅰ ONE，Promega）	1 μl/μg 起始 mRNA 或总 RNA

在 37 ℃下孵育 30 min。

2. 将样品置于冰上，并加入 4 μl SDS〈!〉(10%）和 3 μl 蛋白酶 K（10 μg/μl）以终止 RNase Ⅰ反应。在 45 ℃下孵育 15 min。
3. 用苯酚〈!〉抽提终止蛋白酶 K 反应。将上清液转移到洁净的试管中，并用 H_2O 反萃取使用过的苯酚。合并水相，用 Microcon YM-100 微浓缩装置过滤。

   作为替代，在蛋白酶 K 处理后，用酚-氯仿〈!〉抽提，再用氯仿抽提。用异丙醇和20 μg tRNA 作为载体，反萃取并沉淀核酸。

4. 将团块溶解在 20 μl 0.1×TE 中。

656

磁珠的预处理

5. 为每个文库用 100 μg 不含 DNA 的 tRNA 预处理 MPG 微珠（CPG Inc.，New Jersey)（每 25 μg 起始 mRNA 需要 500 μl)。在冰上孵育 30 min，间或混合。
6. 将试管竖直放置在磁性台座（由磁珠制造商提供）上 3 min，分离磁珠。小心弃

去上清液，并用结合缓冲液（4.5 mol/L NaCl，50 mmol/L EDTA，pH 8.0，每次清洗使用500 μl）清洗3次。

俘获全长的cDNA

7. 将用tRNA处理过的MPG磁珠重悬在500 μl清洗/结合缓冲液中，并将350 μl重悬溶液转移到含有步骤4得到的生物素化第一链cDNA（20 μl cDNA溶解在0.1×TE中）的新试管中。轻轻混匀，在50 ℃下将反应孵育10 min，并直立试管旋转混合。

8. 孵育结束后，将剩下的150 μl微珠加入到反应管中，并在50 ℃继续孵育20 min。

9. 将试管竖直放置在磁性台座上3 min，分离磁珠。

10. 小心弃去上清液，并按如下的步骤清洗磁珠（洗涤体积是0.5 ml）：

    2×清洗/结合溶液（4.5 mol/L NaCl，50 mmol/L EDTA，pH 8.0）

    1×0.3 mol/L NaCl，1 mmol/L EDTA

    2×0.4% SDS，0.5 mol/L NaAcO，20 mmol/L Tris-HCl（pH 8.5），1 mmol/L EDTA

    2×0.5 mol/L NaAcO，10 mmol/L Tris-HCl（pH 8.5），1 mmol/L EDTA

11. （可选对照）移取50 μl未被俘获的cDNA溶液（步骤9的上清液），并用9倍体积的水稀释。用0.6倍体积的异丙醇和2 μg糖原沉淀核酸，置于旁边，用来和下面被俘获的那部分cDNA进行比较（见步骤16)。

12. 最后一次清洗后，加入100 μl 50 mmol/L NaOH、5 mmol/L EDTA以洗脱cDNA。短暂搅动，并在室温下孵育5 min，或混合。

13. 分离磁珠，将上清液（含有洗脱的cDNA）转移到新试管中。将cDNA保存在冰上。

14. 使用100 μl 50 mmol/L NaOH、5 mmol/L EDTA重复洗脱2次。检测洗脱液和磁珠上的计数，检查洗脱是否完全（80%～90%）。

15. 合并洗脱液并保存在冰上，尤其是在用Tris中和以后，以防止cDNA与任何可能存在的RNA污染之间的杂交。这对均一化过程而言非常重要。

16. 移取10 μl洗脱液，与上面步骤10中未被俘获cDNA一起用碱性琼脂糖凝胶进行分析。未结合的那部分cDNA应该比被俘获的那部分要稍微短一些。

17. 将cDNA保存在冰上，并继续加入5′端引导位点的操作（阶段Ⅸ）。

阶段Ⅸ：向cDNA中加入一个5′端引导位点

替代方案：单链接头连接法（SSLM）

有很多方法可以用来实现引导第二链合成的位点（关于完整的讨论，见Sambrook and Russell 2001第11章的导言）。向单链cDNA产物的末端加入双链的DNA接头来

引导第二链的合成，提供了一种可以替代传统加尾的方法（加入单链的引物）。本方案中介绍了 cDNA 的 5′端与 DNA 接头之间的连接（Shibata et al. 2001）。

材料

注意： 标记有〈!〉的材料的正确处理方法参见附录 3。

寡核苷酸引物

准备下列寡核苷酸，每种当中都含有一个 *Xho* Ⅰ限制性位点：

- 寡核苷酸 GN5：
  5′-AGAGAGAGAGCTCGAGCTCTATTTAGGTGACACTATAGAACCAGNNNNN-3′
- 寡核苷酸 N6：
  5′-AGAGAGAGAGCTCGAGCTCTATTTAGGTGACACTATAGAACCANNNNNN-3′
- 寡核苷酸“down”：
  5′P-TGGTTCTATAGTGTCACCTAAATAGAGCTCGAGCTCTCTCTCT- NH_2-3′

P 表明寡核苷酸必须是 5′-磷酸化的，N 表示加入了一个氨基连接以避免非特异性的连接和可能的发夹结构自引导。使用聚丙烯酰胺凝胶电泳纯化寡核苷酸（Sambrook and Russell 2001），并依次用酚：氯仿〈!〉和氯仿抽提。

TaKaRa DNA Ligation Kit Ⅱ连接反应试剂盒（TaKaRa）

方法

1. 将装有合并的第一链产物洗脱液（阶段 X，步骤 15）的试管保持在冰上，并加入 50 μl 1mol/L Tris-HCl（pH 7.0）和 1 μl RNase Ⅰ（10μg/μl RNase ONE，Promega）。迅速混合，在 37 ℃下孵育 10 min。
2. 用蛋白酶 K 处理终止 RNase Ⅰ反应，接着用酚：氯仿抽提，再反萃取一次。
3. 加入 3 μg 糖原，用 Microcon YM-100 微量浓缩装置过滤样品。
4. 根据制造商的说明，用 S400（Amersham Biosciences）或者相当的离心层析柱分离 cDNA，除去痕量的第一链引物。将 cDNA 上样到柱上，然后加入 80 μl H_2O。
5. 按照制造商的说明继续实验，用 Centricon YM-100 微量浓缩装置或者异丙醇沉淀浓缩最终的样品。
6. 将最终样品溶解在 10 μl 0.1×TE 中。
7. 分别用 2.5 倍体积的乙醇沉淀三种寡核苷酸，用 80%乙醇清洗，最后分别溶解在 H_2O 中，浓度为 2 μg/μl。
8. 检查寡核苷酸引物溶液的 OD_{260}，并按照 4：1：5 的比例分别将 GN5、N6 和“down”寡核苷酸混合。
9. 加入 NaCl 至最终浓度为 100 mmol/L，并加热混合物在 65℃下 5 min、45℃下 5 min、37℃下 10 min 和 25℃下 10 min 以进行复性。复性后的接头混合物可以在 −20℃下长期保存。

658

第一链 cDNA 的连接反应

10. 将 cDNA 在 65 ℃下孵育 5 min，使得单链 cDNA 的任何二级结构变性，将试管

转移到冰上。

11. 使用 TaKaRa DNA Ligation Kit Ⅱ建立接头：cDNA 的连接反应。2 μg 接头混合物最多可供 1 μg 单链 cDNA 使用。处理大量的 cDNA 时，要按比例增加接头混合物的用量。
12. 将复性的接头混合物和 cDNA 混合，终体积为 5 μl。
13. 加入 5 μl 来自 TaKaRa DNA Ligation Kit 中的溶液Ⅱ，然后加入 10 μl 溶液Ⅰ。
14. 将反应在 10 ℃下孵育过夜（至少 10 h）。
15. 加入下列试剂终止连接反应：

0.5 mol/L EDTA	1 μl
10% SDS	1 μl
10 mg/ml 蛋白酶 K	1 μl
H_2O	10 μl

将反应在 45 ℃下孵育 15 min。

16. 依次用酚：氯仿〈!〉和氯仿抽提，终止蛋白酶 K 的反应。用 60 μl 柱床缓冲液（10 mmol/L Tris，pH 8.0；1 mmol/L EDTA；0.1% SDS；100 mmol/L NaCl）反萃取。
17. 连接反应结束后，使用 S300 离心层析柱进行纯化，除去过量的接头。
    a. 按照制造商的说明准备纯化柱，加入 2 ml 柱床缓冲液。让柱体在重力作用下排出缓冲液，重复该步骤。
    b. 向柱上加入 100 μl 缓冲液并离心。检查流出洗脱液的体积应为 100 μl。将样品（步骤 16）上柱，以 400 g 的速度离心 2 min。

       如果洗脱液的体积不是 100 μl，按照步骤 a 的方法重新平衡柱床。

    c. 洗脱样品，将滤出液体按照每 50 μl 体积分步收集。
18. 对分步收集的样品进行闪烁计数（见 Sambrook and Russell 2001，附录 8）。合并分步洗脱液中最初的 2 或 3 个分离峰，或者是那些包含了 80%计数的分布洗脱液。
19. 通过异丙醇沉淀（在有 0.2 mol/L NaCl 存在时）或者是 Microcon YM-100 微量浓缩装置浓缩样品。

阶段Ⅺ：用于 SSLLM cDNA 的第二链 cDNA 合成

659

以 SSLLM 和（或）均一化及扣除处理过的第一链 cDNA 作为第二链 cDNA 合成的底物。第二链通过使用长链的热稳定聚合酶混合物（Cheng et al. 1994）进行引物延伸而得到，这样可以高效地合成长链的第二链 cDNA。推荐使用 Invitrogen 的 Elongase 混合酶。

与第一链 cDNA 中的情况相同，第二链 cDNA 的引物-接头中应该包含有复性位点和一个限制性位点，如 *Xho* Ⅰ。对一小部分反应进行放射性标记，以便于监测产量和大小。由于 *Taq* 聚合酶混合物的聚合速度很快，因此混合操作要非常迅速。

方法

1. 合成适当的第二链引物。对于按照阶段Ⅸ中介绍所处理的 cDNA，可使用下面的引物：

   5′-AGAGAGAGAGCTCGAGCTCTATTTAGGTGACACTATAGAACCA-3′

   该引物中包含了一个 *Xho* Ⅰ位点和一个（GA）$_7$ 重复片段用来保证高效的限制性消化。

2. 准备包含了下列试剂的试管 A 和试管 B：

   试管 A：

第一链 cDNA	
100 ng/μl 第二链引物接头	6 μl
5×eLONGase A 缓冲液（Invitrogen）	5.2 μl
5×eLONGase B 缓冲液（Invitrogen）	4.8 μl
2.5 mmol/L（每种）dNTP	6 μl

   用水将最终体积调整到 45 μl。

   试管 B：

[α-^{32}P] dGTP	0.5 μl

3. 第二链合成的热循环仪按照如下方式设定程序：

   步骤 1：65 ℃下 5 min

   步骤 2：68 ℃下 30 min

   步骤 3：72 ℃下 10 min

   步骤 4：+4 ℃

4. 将试管 A 和 B 转移到热循环仪中，启动第二链的程序。当加热块达到 65℃时，向试管 A 中加入 1.5 μl eLONGase 混合物（1 单位/μl）。
5. 将试管中的反应物快速而完全地混合，然后迅速转移 3～5 μl 其中的溶液到试管 B 中。
6. 继续进行程序，然后将试管转移到冰上。
7. 使用 DE-81 吸收定量第二链反应的产量（见 Sambrook and Russell 2001，附录 8）。

   根据所用方案的不同，产量的变化很大；但是，假使 cDNA 是高质量的，那么在本阶段即使是低到 20～30 ng 产量的均一化 cDNA 已经足够用来制备代表性的 cDNA 文库。

8. 用蛋白酶 K 处理第二链 cDNA 产物，用酚-氯仿抽提，反萃取，并用乙醇或者是异丙醇沉淀。

参考文献

Carninci P. and Hayashizaki Y. 1999. High efficiency full-length cDNA cloning. *Methods Ezymol.* **303:** 19–44.

Carninci P., Nakamura M., Sato K., Hayashizaki K., and Brownstein M.J. 2002. Cytoplasmic RNA extraction from fresh and frozen mammalian tissues. *BioTechniques* **33:** 306–309.

Carninci P., Nishiyama Y., Westover A., Itoh M., Nagaoka S., Sasaki N., Okazaki Y., Muramatsu M., and Hayashizaki Y. 1998. Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. *Proc. Natl. Acad. Sci.* **95:** 520–524.

Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., and Hayashizaki Y. 2000. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* **10:** 1617–1630.

Carninci P., Shibata Y., Hayatsu N., Itoh M., Shiraki T., Hirozane T., Watahiki A., Shibata K., Konno H., Muramatsu M., and Hayashizaki Y. 2001. Balanced-size and long-size cloning of full-length, cap-trapped cDNAs into vectors of the novel λ-FLC family allows enhanced gene discovery rate and functional analysis. *Genomics* **77:** 79–90.

Cheng S., Fockler C., Barnes W.M., and Higuchi R. 1994. Effective amplification of long targets from cloned inserts and human genomic DNA. *Proc. Natl. Acad. Sci.* **91:** 5695–5699.

Deininger P.L. 1987. Full-length cDNA clones: Vector-primed cDNA synthesis. *Methods Enzymol.* **152:** 371–389.

Hames B.D. and Higgins S.J., eds. 1985. *Nucleic acid hybridization: A practical approach.* IRL Press, Oxford, United Kingdom.

Kotsopoulos S.T. and Shuber A.P. 1996. Isolation of 3.5-kb fragments on magnetic solid supports. *BioTechniques* **20:** 198–200.

Sambrook J. and Russel D.W. 2001. *Molecular cloning: A laboratory manual,* 3rd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Shibata Y., Carninci P., Watahiki A., Shiraki T., Konno H., Muramatsu M., and Hayashizaki Y. 2001. Cloning full-length, cap-trapper-selected cDNAs by using the single-strand linker ligation method. *BioTechniques* **30:** 1250–1253.

附录 2　试剂、缓冲液和培养基

分子生物学中使用的缓冲液和储存液

注意：标记有〈!〉的材料的正确处理方法参见附录3。

pH缓冲液

磷酸盐缓冲液（PBS）

137 mmol/L NaCl

2.7 mmol/L KCl

10 mmol/L Na_2HPO_4

2 mmol/L KH_2PO_4

将8 g NaCl、0.2 g KCl、1.44 g Na_2HPO_4 和0.24 g KH_2PO_4溶解在800 ml蒸馏水中。用HCl将pH调至7.4，加水至体积为1 L。将溶液分装后在15psi（1.05 kg/cm^2）下高压蒸汽灭菌20 min，或者采用过滤除菌。缓冲液在室温下保存。

PBS是一种通用试剂，改进后适用于不同的应用场合。要注意该配方中不含二价阳离子。如果需要，可向PBS中补加1 mmol/L $CaCl_2$ 和0.5 mmol/L $MgCl_2$。

10×Tris EDTA（TE）

pH7.4

100 mmol/L Tris-Cl（pH 7.4）

10 mmol/L EDTA（pH 8.0）

pH 7.6

100 mmol/L Tris-Cl（pH 7.6）

10 mmol/L EDTA（pH 8.0）

pH 8.0

100 mmol/L Tris-Cl（pH 8.0）

10 mmol/L EDTA（pH 8.0）

将溶液在15 psi（1.05 kg/cm^2）下高压蒸汽灭菌20 min。

缓冲液在室温下保存。

Tris-Cl（1mol/L）

将121.1 g Tris碱溶解在800 ml蒸馏水中。加入浓盐酸〈!〉调节至所需的pH。

pH	HCl
7.4	70 ml
7.6	60 ml
8.0	42 ml

最后调定pH之前，让溶液冷却至室温。加水定容至1 L。分装后高压蒸汽灭菌。

如果1 mol/L溶液呈现黄色，应当弃去，另使用质量更好的Tris。Tris溶液的pH因温度而异，温度每上升1 ℃，pH要下降约0.03单位。例如，浓度为0.05 mol/L的溶液在5 ℃、25 ℃和37 ℃时的pH值分别是9.5、8.9和8.6。

663 Tris镁盐缓冲液（TM）

50 mmol/L Tris-Cl（pH 7.8）

10 mmol/L $MgSO_4$

Tris缓冲盐溶液（TBS）

将8 g NaCl、0.2 g KCl和3 g Tris碱溶解在800 ml蒸馏水中。加入0.015 g酚红并用HCl调节pH至7.4，加入蒸馏水定容至1 L。将溶液分装后在15 psi（1.05 kg/cm^2）下高压蒸汽灭菌20 min。缓冲液在室温下保存。

酶储存液和缓冲液

酶储存液

溶菌酶（10 mg/ml）

使用前即刻将固体溶菌酶溶解在10 mmol/L Tris-Cl（pH 8.0）中，配制浓度为10 mg/ml的溶液。溶解蛋白质之前要确保Tris溶液的pH为8.0。如果溶液的pH低于8.0，溶菌酶则不能有效地发挥效能。

胰DNase Ⅰ（1 mg/ml）

将2 mg粗制胰DNase Ⅰ（Sigma或相当的产品）溶解在1 ml下列溶液中：

10 mmol/L Tris-Cl（pH 7.5）

150 mmol/L NaCl

1 mmol/L $MgCl_2$

在DNase Ⅰ溶解后，加入1 ml甘油，盖紧试管轻轻颠倒几次以均匀混合。要避免产生气泡和泡沫。将溶液分装后在−20 ℃下保存。

胰RNase（1 mg/ml）

将2 mg粗制胰RNase Ⅰ（Sigma或相当的产品）溶解在2 ml TE（pH 7.6）中。

蛋白酶K（20 mg/ml）

购买到的蛋白酶K是冻干粉，溶解在灭菌的50 mmol/L Tris（pH 8.0），1.5 mmol/L乙酸钙中，浓度为20 mg/ml。将储存液分装后在−20 ℃下保存。每份分装的溶液使用时可以冻融数次，但此后应该弃去。与许多粗制的蛋白水解酶（如链霉蛋白酶）不同，蛋白酶K使用前是不需要自我消化处理的。

胰蛋白酶

在200 mmol/L碳酸铵（pH 8.9）（测序级纯度；Boehringer Mannheim）溶液中配制浓度为250 μg/ml的小牛胰蛋白酶〈!〉溶液。将溶液分装后在−20 ℃下保存。

酶反应缓冲液

重要事项：尽可能使用酶生产商提供的10×反应缓冲液。否则，使用此处给出的配方。 664

10×扩增缓冲液

500 mmol/L KCl

100 mmol/L Tris-Cl（pH 8.3，室温下）

15 mmol/L $MgCl_2$

将10×缓冲液在15 psi（1.05 kg/cm^2）下高压蒸汽灭菌10 min。将溶液分装后在−20 ℃下保存。

10×噬菌体T4 DNA连接酶缓冲液

200 mmol/L Tris-Cl（pH 7.6）

50 mmol/L $MgCl_2$

50 mmol/L 二硫苏糖醇〈!〉

0.5 mg/ml 牛血清白蛋白（组分V；Sigma）（可选的）

将缓冲液分装后在−20 ℃下保存。当要反应时加入ATP。反应中的ATP要加入至合适的浓度（如1 mmol/L）。

10×噬菌体T4 DNA聚合酶缓冲液

330 mmol/L Tris-乙酸（pH 8.0）

660 mmol/L 醋酸钾

100 mmol/L 乙酸镁

5 mmol/L 二硫苏糖醇〈!〉

1 mg/ml 牛血清白蛋白（组分V；Sigma）

将10×缓冲液分装后在−20 ℃下保存。

10×噬菌体T4多核苷酸激酶缓冲液

700 mmol/L Tris-Cl（pH 7.6）

100 mmol/L $MgCl_2$

50 mmol/L 二硫苏糖醇〈!〉

将10×缓冲液分装后在−20 ℃下保存。

10×去磷酸化缓冲液（用于CIP）

100 mmol/L Tris-Cl（pH 8.3）

10 mmol/L $MgCl_2$

10 mmol/L $ZnCl_2$

10×去磷酸化缓冲液（用于SAP）

200 mmol/L Tris-Cl（pH 8.8）

100 mmol/L $MgCl_2$

10 mmol/L $ZnCl_2$

10×外切核酸酶Ⅲ缓冲液

660 mmol/L Tris-Cl (pH 8.0)

66 mmol/L $MgCl_2$

100 mmol/L β-巯基乙醇〈!〉

β-巯基乙醇在使用前加入。

10×Klenow 缓冲液

0.4 mol/L 磷酸钾 (pH 7.5)

66 mmol/L $MgCl_2$

10 mmol/L β-巯基乙醇〈!〉

10×接头激酶缓冲液

600 mmol/L Tris-Cl (pH 8.0)

100 mmol/L $MgCl_2$

100 mmol/L 二硫苏糖醇〈!〉

2 mg/ml 牛血清白蛋白

该缓冲液在使用前新鲜配制。

10×蛋白酶 K 缓冲液

100 mmol/L Tris-Cl (pH 8.0)

50 mmol/L EDTA (pH 8.0)

500 mmol/L NaCl

10×反转录酶缓冲液

500 mmol/L Tris-Cl (pH 8.3)

750 mmol/L KCl

30 mmol/L $MgCl_2$

RNase H 缓冲液

20 mmol/L Tris-Cl (pH 7.6)

20 mmol/L KCl

0.1 mmol/L EDTA (pH 8.0)

0.1 mmol/L 二硫苏糖醇〈!〉

该缓冲液在使用前新鲜配制。

5×末端转移酶缓冲液

大多数厂商都提供5×反应缓冲液，其中通常含有：

500 mmol/L 二甲胂酸钾（pH 7.2）〈!〉

10 mmol/L $CoCl_2 \cdot 6H_2O$

1 mmol/L 二硫苏糖醇〈!〉

5×末端转移酶（或加尾）缓冲液可以按照下列方法（Eschenfeldt et al. 1987）来配制：

666

1. 用 10 ml 3 mol/L 醋酸钾在室温下平衡 5 g Chelex 100（Bio-Rad）。
2. 5 min 后，用真空抽吸除去多余的液体。用 10 ml 去离子水清洗 Chelex 3 次。
3. 配制 1 mol/L 二甲胂酸钾溶液。用处理过的 Chelex 树脂平衡二甲胂酸盐溶液。
4. 在布氏漏斗中垫上 Whatman 1 号滤纸，将上一步骤中的二甲胂酸盐溶液进行过滤回收。
5. 向回收的二甲胂酸盐溶液中依次加入水、二硫苏糖醇和氯化钴，配制成最终浓度为 500 mmol/L 二甲胂酸钾，1 mmol/L 二硫苏糖醇和 20 mmol/L $CoCl_2$ 的溶液。将缓冲液分装后在－20 ℃下保存。

10×通用 KGB（限制性内切核酸酶）缓冲液

1 mol/L 醋酸钾

250 mmol/L Tris-乙酸（pH 7.6）

100 mmol/L 四水合乙酸镁

5 mmol/L β-巯基乙醇〈!〉

0.1 mg/ml 牛血清白蛋白

将缓冲液分装后在－20 ℃下保存。

杂交缓冲液

HCl（2.5 mol/L）

向 91 ml 灭菌水中加入 25 ml 浓 HCl〈!〉（11.6 mol/L）。稀释后的溶液在室温下保存。

含甲酰胺的杂交缓冲液（用于 RNA）

40 mmol/L PIPES（pH6.8）

1 mmol/L EDTA（pH 8.0）

0.4 mol/L NaCl

80%（*V/V*）去离子甲酰胺〈!〉

用 PIPES 的二钠盐配制缓冲液，并用 1 mol/L HCl 调节 pH 至 6.4。

不含甲酰胺的杂交缓冲液（用于 RNA）

40 mmol/L PIPES（pH 6.4）

1 mmol/L EDTA (pH 8.0)

0.4 mol/L NaCl

用 PIPES 的二钠盐配制缓冲液，并用 1 mol/L HCl 调节 pH 至 6.4。

预杂交和杂交液

预杂交/杂交液（适用于液体缓冲液中的杂交）

6×SSC（或 6×SSPE）

5×Denhardt 试剂

0.5%（*m/V*）SDS〈!〉

1 μg/ml poly（A）

100 μg/ml 鲑鱼精 DNA

667 预杂交/杂交液（适用于甲酰胺缓冲液中的杂交）

6×SSC（或 6×SSPE）

5×Denhardt 试剂

0.5%（*m/V*）SDS〈!〉

1 μg/ml poly（A）

100 μg/ml 鲑鱼精 DNA

50%（*V/V*）甲酰胺〈!〉

完全混合后，用 0.45 μm 的一次性醋酸纤维素膜（Scheicher & Schuell Uniflow 注射器滤膜或其他同类产品）过滤。为降低在非严格条件下（如 20%～30%甲酰胺）的杂交背景，很重要的是所使用甲酰胺的纯度要尽可能的高。

预杂交/杂交液（适用于磷酸盐-SDS 缓冲液中的杂交）

0.5 mol/L 磷酸盐缓冲液（pH 7.2）*

1 mmol/L EDTA (pH 8.0)

7%（*m/V*）SDS〈!〉

1%（*m/V*）牛血清白蛋白

使用电泳级的牛血清白蛋白。应用这种特殊的预杂交/杂交液不需要封闭试剂或杂交率增强剂。

20×SSC

将 175.3 g NaCl 和 88.2 g 柠檬酸钠溶解在 800 ml 水中。用数滴 14 mol/L 的 HCl〈!〉将 pH 调节至 7.0。用水定容至 1 L。分装后高压蒸汽灭菌。缓冲液的最终浓度为 3.0 mol/L NaCl 和 0.3 mol/L 柠檬酸钠。

20×SSPE

将 175.3 g NaCl，27.6 g $NaH_2PO_4 \cdot H_2O$ 和 7.4 g EDTA 溶解在 800 ml 水中。用

* 0.5 mol/L 磷酸盐缓冲液为 134 g $Na_2HPO_4 \cdot 7H_2O$，4 ml 85% H_3PO_4〈!〉（浓磷酸），加水至 1 L。

NaOH〈!〉（约 6.5 ml 10 mol/L 的溶液）将 pH 调节至 7.4。用水定容至 1 L。分装后高压蒸汽灭菌。缓冲液的最终浓度为 3.0 mol/L NaCl、0.2 mol/L NaH_2PO_4、0.02 mol/L EDTA。

凝胶上样缓冲液

表 A2-1　6×凝胶加样缓冲液

缓冲液类型	6×缓冲液	储存温度
Ⅰ	0.25%（*m/V*）溴酚蓝〈!〉 0.25%（*m/V*）二甲苯青 FF〈!〉 40%（*m/V*）蔗糖水溶液	4 ℃
Ⅱ	0.25%（*m/V*）溴酚蓝〈!〉 0.25%（*m/V*）二甲苯青 FF〈!〉 15%（*m/V*）聚蔗糖（Ficoll）（400 型；Pharmacia 公司）水溶液	室温
Ⅲ	0.25%（*m/V*）溴酚蓝〈!〉 0.25%（*m/V*）二甲苯青 FF〈!〉 30%（*V/V*）甘油水溶液	4 ℃
Ⅳ	0.25%（*m/V*）溴酚蓝〈!〉 40%（*m/V*）蔗糖水溶液	4 ℃

特殊的缓冲液和溶液

STE

10 ml Tris-Cl（pH 8.0）
0.1 mol/L NaCl
1 mmol/L EDTA（pH 8.0）

在 15 psi（1.05 kg/cm^2）下高压蒸汽灭菌 15 min。灭菌后的溶液在 4℃下保存。

STET

10 mmol/L Tris-Cl（pH 8.0）
0.1 mol/L NaCl
1 mmol/L EDTA（pH 8.0）
5%（*V/V*）Triton X-100

所有成分加入后要确保 STET 的 pH 为 8.0。用前不需要对 STET 进行灭菌处理。

10×TEN 缓冲液

0.1 mol/L Tris-Cl（pH 8.0）
0.01 mol/L EDTA（pH 8.0）
1 mol/L NaCl

TES

10 mmol/L Tris-Cl（pH 7.5）

1 mmol/L EDTA (pH 7.5)

0.1% (m/V) SDS 〈!〉

669 Triton/SDS 溶液

10 mmol/L Tris-Cl (pH 8.0)

2% (V/V) Triton X-100

1% (m/V) SDS 〈!〉

100 mmol/L NaCl

1 mmol/L EDTA (pH 8.0)

溶液通过 0.22 μm 滤膜过滤器除菌，在室温下保存。

Tris-蔗糖

50 mmol/L Tris-Cl (pH 8.0)

10% (m/V) 蔗糖

溶液通过 0.22 μm 滤膜过滤器除菌，在室温下保存。含有蔗糖的溶液不应高压蒸汽灭菌，因为糖在高温下容易碳化。

清洗缓冲液 (QIAGEN)

50 mmol/L MOPS-KOH 〈!〉 (pH 7.5~7.6)

0.75 mol/L NaCl

15% (V/V) 乙醇

制备该缓冲液时，在加入乙醇之前要调节 MOPS/NaCl 溶液的 pH。

有机试剂的配制

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

酚

大多数批次的市售液化酚〈!〉是清亮和无色的，无须重蒸即可用于分子克隆。偶尔有些批次的液化酚为粉红或黄色，这种产品应该拒绝接受，并退还生产厂商。不推荐使用结晶酚，因为它必须在 160 ℃进行重蒸，以去除氧化产物，如醌等，这些产物会使得磷酸二酯键断裂，或导致 RNA 和 DNA 的交联。

酚的平衡

酚在使用前必须将其平衡至 pH>7.8，因为 DNA 在酸性条件下会分配入有机相。进行本步骤操作时，要戴上手套，防护面具和实验服。

1. 液化酚储存在−20 ℃下。用前从冰箱中取出，使其温度升至室温，然后在 68 ℃下使其熔化。加入羟基喹啉至最终浓度为 0.1%。羟基喹啉是一种抗氧剂，RNase 的不完全抑制剂及金属离子的弱螯合剂 (Kirby 1956)。此外，其黄色提

供了一种快速识别有机相的方法。

670

2. 向熔化的酚中，加入等体积的缓冲液（通常为 0.5 mol/L Tris-Cl，pH 8.0，室温下）。在磁力搅拌器上将混合物搅拌 15 min。关闭搅拌器，待两相分离后，使用与装有抽滤瓶的真空装置相连的玻璃吸管尽可能多的吸出上层（水）相。
3. 加入等体积的 0.1 mol/L Tris-Cl（pH 8.0）到酚中，在磁力搅拌器上搅拌 15 min。关闭搅拌器，按照步骤 2 所述除去上层水相。重复萃取操作，直到酚相的 pH>7.8（用 pH 试纸测量）。
4. 酚达到平衡，并除去最终的水相后，加入 0.1 体积含有 0.2%β-巯基乙醇〈!〉的 0.1 mol/L Tris-Cl（pH 8.0）。在 4 ℃下，这种形式的酚溶液可以在避光的瓶中，处在 100 mmol/L Tris-Cl（pH 8.0）之下，保存时间长达 1 个月。

酚：氯仿：异戊醇（25：24：1）

等体积的平衡酚和氯仿：异戊醇〈!〉（24：1）组成的混合溶液常用于在核酸制备时除去蛋白质。氯仿的作用是使蛋白质变性，并有助于水相和有机相的分离，异戊醇则可以减少萃取中泡沫的产生。无论氯仿，还是异戊醇在用前都不需要处理。在 4 ℃下，酚：氯仿：异戊醇混合物可以在避光的瓶中，处在 100 mmol/L Tris-Cl（pH 8.0）之下，保存时间长达 1 个月。

甲酰胺的去离子化

许多批次的试剂级甲酰胺〈!〉纯度已经足够，不需要进一步的处理。但是，如果出现黄色，就要加入 Dowex XG8 混合床树脂，在磁力搅拌器上搅拌 1 h，再用 Whatman 1 号滤纸过滤 2 次，以作去离子处理。去离子甲酰胺分装后，在氮气氛围中，于 −70 ℃下保存。

化学试剂储存液

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

丙烯酰胺溶液（45%，*m/V*）

丙烯酰胺（DNA 测序级）〈!〉	434 g
N,*N*-亚甲基双丙烯酰胺〈!〉	16 g
H_2O	定容至 600 ml

将溶液加热到 37 ℃使试剂溶解。用蒸馏水定容至 1L。用硝酸纤维素滤膜（如 Nalge，0.45 μm 孔径）过滤，将过滤后的溶液储存在深色瓶中，在室温下保存。

671

乙酸铵（10 mol/L）

配制 1 L 溶液时，将 770 g 乙酸铵溶解在 800 ml 水中。加水定容至 1 L。过滤除菌。或者，要配制 100 ml 溶液，室温下将 77 g 乙酸铵溶解于 70 ml 水中。加水定容至 100 ml。用 0.22 μm 滤器过滤除菌。溶液装在密封的瓶中，在 4 ℃或室温下保存。乙酸铵在热水中分解，所以含有乙酸铵的溶液不能用高压蒸汽灭菌。

过硫酸铵（10%，*m/V*）

过硫酸铵〈!〉	1 g
H_2O	定容到 10 ml

在 10 ml 水中溶解 1 g 过硫酸铵，4 ℃下储存。过硫酸铵在溶液中会缓慢分解，因此每 2～3 周要更换储存液。过硫酸铵用作丙烯酰胺和亚甲基双丙烯酰胺共聚合的催化剂。聚合反应是由氧化-还原反应产生的自由基驱动的，在氧化-还原反应中，二胺（如 TEMED）被用作附加催化剂（Chrambach and Rodbard 1972）。

ATP（10 mol/L）

在 25 mmol/L Tris-Cl（pH8.0）中溶解适量的固态 ATP 配制。ATP 溶液分装后在－20 ℃下保存。

氯化钙（2.5 mol/L）

将 11 g $CaCl_2 \cdot 6H_2O$ 溶解在蒸馏水中，定容到 20 ml。用 0.22 μm 的滤器过滤除菌，溶液按 1 ml 分装后在 4 ℃下保存。

含有四种 dNTP 的 dNTP 溶液，每种的浓度均为 10 mmol/L

储存液分装后在－20℃下保存，使用时可进行稀释，如稀释成含有全部四种 dNTP 的、浓度均为 2.5 mmol/L 的样品。使用微量天平称取所需用量的每种脱氧核苷三磷酸，并装入灭菌的微量离心管中。配制不同 dNTP 溶液时，每次称量中使用一次性的刮铲，或者在每次称量之间用乙醇仔细清洁刮铲。下表列出了配制 1ml 浓度10 mmol/L 储存液时所需的固态无水 dNTP 的用量。

dNTP	FW	配制 1 ml 浓度为 10 mmol/L 溶液时需要的用量/mg
dATP	491.2	4.91
dCTP	467.3	4.67
dTTP	482.2	4.82
dGTP	507.2	5.07

672

将脱氧核苷三磷酸溶解在少量的水中，然后用 pH 试纸和自动移液装置，加入少量 2 mol/L 的 NaOH 将 pH 调节至约 8.0。或者，也可以使用脱氧核苷三磷酸的单钠盐来配制储存液，此时要将适量的固体溶解在 1 ml TE（pH 8.0）中。将管中的试剂混匀，储存液分装后在－20 ℃下保存。

许多生产厂商供应调整好 pH 的 dNTP 溶液。

二甲基亚砜（DMSO）

购买高试剂等级的 DMSO〈!〉（HPLC 级或者更好的）。将新开瓶的试剂按照 1 ml

分装到灭菌的试管中。密封试管，在－20 ℃下保存。每份分装的试剂使用一次后弃去。

二硫苏糖醇（DTT，1mol/L）

在 20 ml 0.01mol/L 醋酸钠（pH 5.2）中溶解 3.09 g 二硫苏糖醇〈!〉，过滤除菌。将溶液按照 1 ml 分装，在－20 ℃下保存。在此条件下，二硫苏糖醇保持稳定，不会被空气氧化。

EDTA（0.5 mol/L，pH 8.0）

将 186.1 g 二水合乙二胺四乙酸二钠加入 800 ml 水中。在磁力搅拌器上剧烈搅拌，加入 NaOH（约 20 g 颗粒状 NaOH〈!〉）调节 pH 至 8.0。分装后高压蒸汽灭菌。除非加入 NaOH 调节 pH 至 8.0，否则 EDTA 二钠盐不会溶解。

EGTA（0.5 mol/L，pH 8.0）

EGTA 是乙二醇双（2-氨乙基醚）四乙酸。EGTA 溶液的配制基本上是按照上述的 EDTA 的方法配制，可以使用高压蒸汽灭菌，或者过滤除菌。溶液灭菌后在室温下保存。

溴化乙啶（10 mg/ml）

将 1 g 溴化乙啶〈!〉溶解在 100 ml 水中。在磁力搅拌器上搅拌数小时，确保染料溶解。用铝箔将容器包裹起来，或者将溶液转移到深色瓶中，在室温下保存。

明胶（2%，*m/V*）

将 2 g 凝胶加入到总体积为 100 ml 的水中，溶液在 15 psi（1.05 kg/cm^2）下高压蒸汽灭菌 15min。

甘油（10%，*V/V*）

将 1 体积的分子生物学级甘油用 9 体积的灭菌纯水稀释。溶液用事先清洗过的 0.22 μm 滤器过滤除菌。按照 200 ml 分装后在 4 ℃下保存。

KCl（4 mol/L）

将适量的固态氯化钾溶解在水中，高压蒸汽灭菌 20 min，在室温下保存。理想的做法是将溶液分装成小份（约 100 μl）后保存在无菌试管中，并且每份只供给一次使用。

线性聚丙烯酰胺

线性聚丙烯酰胺是使用乙醇沉淀皮克级核酸时的一种高效中性载体（Gaillard and Strauss 1990）。制备线性聚丙烯酰胺时，在 40 mmol/L Tris-Cl（pH 8.3）、20 mmol/L 醋酸钠、1 mmol/L EDTA（pH 8.0）中配制 5%的丙烯酰胺溶液（不含双丙烯酰胺）。加入 1/100 体积新鲜配制的 10%过硫酸铵和 1/1000 体积的 TEMED。当溶液变黏稠后，

用 2.5 倍体积的乙醇沉淀聚合物。在 20 体积的水中振荡过夜溶解聚合物团块。0.25%的线性聚丙烯酰胺溶液在 4 ℃下可以稳定保存数年。

氯化锂（LiCl，5mol/L）

在最终体积为 90 ml 的水中溶解 21.2 g LiCl。用水定容到 100 ml。溶液通过 0.22 μm的滤器过滤除菌，或者在 15 psi（1.05 kg/cm^2）下高压蒸汽灭菌 15 min。在 4 ℃下保存溶液。

$MgCl_2 \cdot 6H_2O$（1 mol/L）

在 800 ml 水中溶解 203.3 g $MgCl_2 \cdot 6H_2O$，用水定容到 1 L。分装后高压蒸汽灭菌。氯化镁极易潮解，应选购小瓶（如 100 g）试剂，启用后勿长期存放。

$MgSO_4$（1 mol/L）

在最终体积为 100 ml 的水中溶解 12 g $MgSO_4$。高压蒸汽灭菌或过滤除菌。在室温下保存。

NaOH（10 mol/L）

10 mol/L NaOH〈!〉的配制是一个剧烈放热反应，可能会造成玻璃容器的破裂。要特别小心地在塑料烧杯中配制这种溶液。向 800 ml 水中缓慢加入 400 g NaOH 颗粒〈!〉，持续搅拌。作为预防措施，将烧杯置于冰上。当颗粒完全溶解后，用水定容至 1 L。溶液装在塑料容器中，于室温下保存。不需要灭菌。

NaCl（5 mol/L）

将 292 g NaCl 溶解在 800 ml 水中，用水定容至 1 L。分装后高压蒸汽灭菌。NaCl 溶液在室温下保存。

含有全部四种核糖核苷三磷酸的 NTP 溶液，每种浓度均为 2.5 mmol/L

储存液可以用核糖核苷三磷酸的三钠盐来配制，将适量的固体溶解在 1.0 ml 的 20 mmol/L Tris-Cl（pH 7.6）中。混匀试管内物质，储存液分装后在 −20 ℃下保存。

核糖核苷酸	FW	配制 1 ml 浓度为 2.5 mmol/L 溶液时需要的量/mg
ATP-3Na	573.1	1.43
CTP-3Na	549.1	1.37
TTP-3Na	589.2	1.47
GTP-3Na	550.1	1.37

注：许多生产厂商供应 pH 调整好的 NTP 溶液。

PEG 8000

PEG〈!〉的工作浓度为13%～40%（*m/V*）。用灭菌水溶解 PEG 8000，配制成适当浓度的溶液，如果需要可以加热。用0.22 μm 滤器过滤除菌。溶液在室温下保存。

> 聚乙二醇（PEG）是一种单一重复单位 H（OCH_2CH_2）$_n$OH 的直链聚合物。PEG 具有多种分子质量，其命名反映了每种分子中重复单元的数量（*n*）。例如，在 PEG400 中，*n*=8 或 *n*=9，而在 PEG4000 中，*n* 的范围为 68～84。PEG 在水溶液中诱导溶质大分子的凝聚（Zimmerman and Minton 1993），在分子克隆中有着多种用途，包括：
>
> • 根据分子质量的大小沉淀 DNA。沉淀所需的 PEG 浓度与 DNA 片段的大小成反比。（Lis and Schleif 1975a，b；Ogata and Gilbert 1977；Lis 1980）。
>
> • 噬菌体颗粒的沉淀和纯化（Yamamoto et al.）。
>
> • 在杂交，DNA 分子的平端连接和用噬菌体 T4 多核苷酸激酶对中 DNA 进行末端标记中提高核酸互补链的复合效率。
>
> • 培养细胞与细菌原生质体的融合（Schaffner 1980；Rassoulzadegan et al. 1982）

醋酸钾（5 mol/L）

5 mol/L 乙酸钾	60 ml
冰乙酸〈!〉	11.5 ml
H_2O	28.5 ml

配制的溶液中钾的浓度为 3 mol/L，乙酸的浓度为 5 mol/L。该缓冲液在室温下保存。

SDS（20%，*m/V*）

亦称作十二烷基磺酸钠。在 900 ml 水中溶解 200 g 电泳级的 SDS〈!〉。加热至68 ℃，并用磁力搅拌器搅拌可促进溶解。如果需要，可加入数滴浓 HCl〈!〉将 pH 调节至7.2。用水定容至1 L。室温下保存，无须灭菌。不可高压蒸汽灭菌。

醋酸钠（3 mol/L，pH 5.2 和 pH 7.0）

将 408.3 g 三水合乙酸钠溶解在 800 ml 水中，用冰乙酸〈!〉调节 pH 至5.2，或用稀乙酸调节 pH 至7.0。用水定容至1 L。分装后高压蒸汽灭菌。

亚精胺（1 mol/L）

将 1.45 g 亚精胺（游离碱形式）溶解在 10 ml 去离子水中，并用 0.22 μm 滤器过滤除菌。分装后在−20 ℃下保存。该试剂每个月都要求重新配制。

三氯乙酸（TCA，100%溶液）

向新开启的瓶装 500 g TCA〈!〉中加入 227 ml 水，配制的溶液中包含 100%（*m/V*）TCA。

大肠杆菌液体培养基

注意：标记有〈!〉的材料的正确处理方法参见附录3。

GYT 培养基（Tung and Chow 1995）

10%（*V*/*V*）甘油

0.125%（*m*/*V*）酵母提取物

0.25%（*m*/*V*）胰化蛋白胨

使用预先冲洗过的 0.22 μm 滤器过滤除菌。按 2.5 ml 分装后在 4 ℃下保存。

LB（Luria-Bertani）培养基

配制每升培养基时：

向 950 ml 去离子水中加入：

胰化蛋白胨	10 g
酵母提取物	5 g
NaCl	10 g

摇动容器直到溶质全部溶解。用 5 mol/L NaOH（约 0.2 ml）调节 pH 至 7.0。用去离子定容至 1 L。在 15 psi（1.05 kg/cm^2）下高压蒸汽灭菌 20 min。

M9 基本培养基

配制每升培养基时：

向 750 ml 无菌水（冷却至 50 ℃或更低温度）中加入：

5×M9 盐溶液*	200 ml
1 mol/L $MgSO_4$	2 ml
合适碳源的 20%溶液（如 20%葡萄糖）	20 ml
1 mol/L $CaCl_2$	0.1 ml
灭菌去离子水	定容至 980 ml

如果需要，向 M9 培养基内补加适当的氨基酸和维生素的储存液。

676

* 在去离子水中溶解下列盐类，最终体积 1L，配制成 5×M9 盐溶液：

$Na_2HPO_4 \cdot 7H_2O$	64 g
KH_2PO_4	15 g
NaCl	2.5 g
NH_4Cl	5.0 g

将盐溶液按照 200 ml 分装，在 15 psi（1.05 kg/cm^2）下高压蒸汽灭菌 15 min。

$MgSO_4$ 和 $CaCl_2$ 溶液要分开配制，并用高压蒸汽灭菌。将 5×M9 盐溶液用无菌水稀释至 980 ml 后，再加入上述两种溶液。葡萄糖溶液加入到稀释的 M9 盐溶液中之前用 0.22 μm 滤器过滤除菌。

当使用含有染色体上脯氨酸生物合成操纵子［Δ（*lac-pro*AB）］的大肠杆菌以及互补的 *pro*AB 基因在 F′质粒上时，在 M9 基本培养基中补加下列成分：

0.4%（*m*/*V*）葡萄糖（右旋糖）

5 mmol/L $MgSO_4 \cdot 7H_2O$

0.01%硫胺

NZCYM 培养基

配制每升培养基时：

向 950 ml 去离子水中加入：

NZ 胺	10 g
NaCl	5 g
酵母提取物	5 g
（水解）酪蛋白氨基酸	1 g
$MgSO_4 \cdot 7H_2O$	2 g

摇动容器直到溶质全部溶解。用 5 mol/L NaOH（约 0.2 ml）调节 pH 至 7.0。用去离子定容至 1L。在 15 psi（1.05 kg/cm²）下高压蒸汽灭菌 20 min。

NZ 胺：酶促酪蛋白水解物（ICN Biochemicals）。NZCYM、NZYM 和 NZM 也可用 BD Biosicences 的脱水培养基。

SOB 培养基

配制每升培养基时：

向 950 ml 去离子水中加入：

胰化蛋白胨	20 g
酵母提取物	5 g
NaCl	0.5 g

摇动容器直到溶质全部溶解。加入 10 ml 250 mmol/L 的 KCl 溶液（将 1.86 g KCl 溶解在 100 ml 去离子水中配制而成）。用 5 mol/L NaOH（约 0.2 ml）将 pH 调节至 7.0。用去离子定容至 1 L。在 15 psi（1.05 kg/cm²）下高压蒸汽灭菌 20min。该溶液使用前加入 5 ml 灭菌的 2mol/L $MgCl_2$ 溶液。[2mol/L $MgCl_2$ 溶液配制方法：将 19 g $MgCl_2$ 溶解在 90 ml 去离子水中。用去离子水定容至 100 ml。在 15 psi（1.05 kg/cm²）下高压蒸汽灭菌 20 min]。

Terrific 肉汤（也称为 TB 培养基；Tartof and Hobbs 1987）

677

配制每升培养基时：

向 900 ml 去离子水中加入：

胰化蛋白胨	12 g
酵母提取物	24 g
甘油	4 ml

摇动容器直到溶质全部溶解，然后在 15 psi（1.05 kg/cm²）下高压蒸汽灭菌 20 min。当溶液冷却至 60 ℃或更低温度时，加入 100 ml 灭菌的 0.17 mol/L KH_2PO_4 和 K_2HPO_4 溶液［配制该溶液时，向 90 ml 去离子水中加入 2.31g KH_2PO_4 和 12.54 g K_2HPO_4，溶解后用去离子水定容至 100 ml，然后在 15 psi（1.05 kg/cm²）下高压蒸汽灭菌 20 min］。

2× YT 培养基

配制每升培养基时：

向 900 ml 去离子水中加入：

胰化蛋白胨	16 g
酵母提取物	10 g
NaCl	5 g

摇动容器直到溶质全部溶解。用 5 mol/L NaOH（约 0.2 ml）调节 pH 至 7.0。用去离子定容至 1 L。在 15 psi（1.05 kg/cm^2）下高压蒸汽灭菌 20 min。

含琼脂或琼脂糖的培养基

注意：标记有〈!〉的材料的正确处理方法参见附录 3。

按照上述配方配制液体培养基。高压蒸汽灭菌前加入下列试剂之一：

细菌培养用琼脂（Bacto Agar）（铺制平板用）	15 g/L
细菌培养用琼脂（Bacto Agar）（配制顶层琼脂用）	7 g/L
琼脂糖（铺制平板用）	15 g/L
琼脂糖（配制顶层琼脂用）	7 g/L

在 15 psi（1.05 kg/cm^2）下高压蒸汽灭菌 20 min。从高压蒸汽灭菌锅中取出培养基时，要轻轻旋动使得熔化的琼脂或琼脂糖均匀分布在整个培养基溶液中。必须小心！此时培养基溶液可能过热，旋动时可能会发生暴沸。待培养基冷却到 50～60 ℃后再加入热不稳定性物质（如抗生素）。为避免产生气泡，混匀培养基时要采用旋动方式。然后可以从烧瓶中直接倒制平板。每个 90 mm 的培养皿需要大约 30～35 ml 培养基。为了从平板上的培养基中赶去气泡，在琼脂或琼脂糖凝固之前，用本生灯烧灼培养基表面使气泡散去。设定颜色记号（如两条红杠表示 LB-氨苄青霉素平板，一条黑杠表示 LB 平板等），在相应平板的边缘作上标记。

678

当培养基完全凝固后，将平板倒置并在 4℃下储存备用。平板必须在使用前 1～2 h 从冷藏室取出。如果平板是新铺制的，在 37 ℃下孵育时会“出汗”。如果冷凝的液滴落到琼脂/琼脂糖表面，会导致细菌菌落或噬菌体噬斑在平板表面发生扩散，从而增大交叉污染的机会。为避免出现这一问题，可以拭去平板上盖内部的液滴，将平板倒置，在 37 ℃下孵育数小时后再使用。或者，也可以快速地甩一下平板盖来除去上面的液滴。为了使污染的可能性降至最低，除去盖上的液滴时必须将开盖的平板倒置。

抗生素

表 A2-2　常用抗生素溶液

	储存溶液[a]		工作浓度	
	浓　度	保存条件	严谨型质粒	松弛型质粒
氨苄青霉素	50 mg/ml，水溶液	−20 ℃	20 μg/ml	50 μg/ml
羧苄青霉素	50 mg/ml，水溶液	−20 ℃	20 μg/ml	60 μg/ml
氯霉素	34 mg/ml，乙醇溶液	−20 ℃	25 μg/ml	170 μg/ml
卡那霉素	10 mg/ml，水溶液	−20 ℃	10 μg/ml	50 μg/ml
链霉素	10 mg/ml，水溶液	−20 ℃	10 μg/ml	50 μg/ml
四环素[b]	5mg/ml，乙醇溶液	−20 ℃	10 μg/ml	50 μg/ml

注：镁离子是四环素的拮抗剂。对于以四环素为筛选抗性的细菌，必须使用不含镁盐的培养基（如 LB 培养基）。

a 溶于水的抗生素储存液要使用 0.22 μm 的滤菌过滤除菌。

b 溶于乙醇的抗生素无须除菌。溶液应保存在避光容器内。

用于酵母繁殖与筛选的培养基

注意：标记有〈!〉的材料的正确处理方法参见附录3。

重要提示：在所有配方中使用蒸馏去离子水。除非另有说明，培养基和溶液在15 psi（1.05 kg/cm^2）下高压蒸汽灭菌15～20 min。

完全基本培养基（CM）或合成完全培养基（SC）和营养缺陷型培养基

如果要确定菌株生长的必需条件，采用相应的营养缺陷型培养基是非常有用的。培养基补料的干粉是各自分开保存的，可以方便的配制相应的营养缺陷型培养基。CM（或SC）是含有全部生长成分的培养基（也就是说，没有一种营养成分缺陷）。

无氨基酸酵母氮碱（YNB）*	6.7 g
葡萄糖	20 g
Bacto Agar	20 g
营养缺陷混合物	2 g
水	定容至1000 ml

营养缺陷混合物

混合适当的成分，同时减去有关的补料，在密封的容器中混匀。来回倒置转动容器至少15 min；可以加入数粒干净的大理石颗粒帮助固体成分混匀。

腺嘌呤	0.5 g
丙氨酸	2.0 g
精氨酸	2.0 g
天冬酰胺	2.0 g
天冬氨酸	2.0 g
半胱氨酸	2.0 g
谷氨酸盐	2.0 g
谷氨酸	2.0 g
甘氨酸	2.0 g
组氨酸	2.0 g
肌醇	2.0 g
异亮氨酸	2.0 g
亮氨酸	10.0 g
赖氨酸	2.0 g
蛋氨酸	2.0 g
对氨基苯甲酸	0.2 g
苯丙氨酸	2.0 g
脯氨酸	2.0 g

* 市售的无氨基酸酵母氮碱（YNB）分为含硫酸铵与不含硫酸铵两种。本配方中使用的YNB带有硫酸铵。如果YNB包装上标识为不含硫酸铵的，那么只要加1.7 g YNB，另加5 g硫酸铵即可。

丝氨酸	2.0 g
苏氨酸	2.0 g
色氨酸	2.0 g
酪氨酸	2.0 g
尿嘧啶	2.0 g
缬氨酸	2.0 g

680 补充基本培养基（SMM）

SMM 是添加了各种生长补料的合成葡萄糖基本培养基（SD）。这些溶液可以长期存放。为了防止沉淀，一些溶液要在室温下保存，而其他溶液则需冷藏。无论如何，最好使用氨基酸的 HCl 盐。

配制培养基时，向 SD 培养基的配料中加入适当体积的储存液，然后用蒸馏水定容至 1L。高压蒸汽灭菌后分别加入苏氨酸和天冬氨酸溶液。

作为替代，一种更为方便的培养基配制方法是：将少量补料涂布在 SD 平板表面，待补料液充分干燥后再接种酵母菌。

表 A2-3 中给出了储存液的浓度，配制 1L 培养基时所需储存液的体积，涂布 SD 平板时需要的储存液体积，以及每种成分在 SMM 中的最终浓度。

表 A2-3 补加基本培养基的成分

成 分	储存液浓度/(g/100ml)	1L 中所用储存培养基的体积/ml	培养基中的最终浓度/(mg/L)	涂布在平板上需要的储存液体积/ml
腺嘌呤硫酸盐	0.2[a]	10	20	0.2
尿嘧啶	0.2[a]	10	20	0.2
L-色氨酸	1	2	20	0.1
L-组氨酸 HCl	1	2	20	0.1
L-精氨酸 LiCl	1	2	20	0.1
L-蛋氨酸	1	2	20	0.1
L-酪氨酸	0.2	15	30	0.2
L-亮氨酸	1	10	100	0.1
L-异亮氨酸	1	3	30	0.1
L-赖氨酸 HCl	1	3	30	0.1
L-苯丙氨酸	1[a]	5	50	0.1
L-谷氨酸	1[a]	10	100	0.2
L-天冬氨酸	1[a,b]	10	100	0.2
L-缬氨酸	3	5	150	0.1
L-苏氨酸	4[a,b]	5	200	0.1
L-丝氨酸	8	5	400	0.1

a 室温下保存。

b 在培养基高压蒸汽灭菌后添加。

合成的葡萄糖基本培养基（SD）

SD是含有盐类、微量元素、氮源（不含氨基酸的酵母氮碱）和葡萄糖的合成基本培养基。

不含氨基酸的酵母氮碱（YNB）*	6.7 g
葡萄糖	20 g
Bacto Agar	20 g
水	1000 ml

YPD（YEPD）培养基

YPD是用于酵母常规生长的复合培养基。

酵母提取物	10 g
蛋白胨	20 g
葡萄糖	20 g
水	定容至1000ml

铺制平板时，在高压蒸汽灭菌前加入20 g Bacto Agar（2%）。

参考文献

Chrambach A. and Rodbard D. 1972. Polymerization of polyacrylamide gels: Efficiency and reproducibility as a function of catalyst concentrations. *Sep. Sci.* **7:** 663–703.

Gaillard C. and Strauss F. 1990. Ethanol precipitation of DNA with linear acrylamide as a carrier. *Nucleic Acids Res.* **18:** 378.

Kirby K.S. 1956. A new method for the isolation of ribonucleic acids from mammalian tissues. *Biochem. J.* **64:** 405.

Lis J.T. 1980. Fractionation of DNA fragments by polyethylene glycol induced precipitation. *Methods Enzymol.* **65:** 347–353.

Lis J.T. and Schleif R. 1975a. Size fractionation of double-stranded DNA by precipitation with polyethylene glycol. *Nucleic Acids Res.* **2:** 383–389.

———. 1975b. The regulatory region of the L-arabinose operon: Its isolation on a 1000 base-pair fragment with DNA heteroduplexes. *J. Mol. Biol.* **95:** 409–416.

Ogata R. and Gilbert W. 1977. Contacts between the *lac* repressor and the thymines in the *lac* operator. *Proc. Natl. Acad. Sci.* **74:** 4973–4976.

Rassoulzadegan M., Binetruy B., and Cuzin F. 1982. High frequency of gene transfer after fusion between bacteria and eukaryotic cells. *Nature* **295:** 257–259.

Schaffner W. 1980. Direct transfer of cloned genes from bacteria to mammalian cells. *Proc. Natl. Acad. Sci.* **77:** 2163–2167.

Tartof K.D. and Hobbs C.A. 1987. Improved media for growing plasmid and cosmid clones. *Focus* (Life Technologies) **9:** 12.

Tung W.L. and Chow K.C. 1995. A modified medium for efficient electrotransformation of *E. coli*. *Trends Genet.* **1:** 128–129.

Yamamoto K.R., Alberts B.M., Benzinger R., Lawhorne L., and Treiber G. 1970. Rapid bacteriophage sedimentation in the presence of polyethylene glycol and its application to large-scale virus purification. *Virology* **40:** 734–744.

Zimmerman S.B. and Minton A.P. 1993. Macromolecular crowding: Biochemical, biophysical, and physiological consequences. *Annu. Rev. Biophys. Biomol. Struct.* **22:** 27–65.

* 见上面CM配方的说明。

683

附录3 注意事项

一般注意事项

无论何时何地都要遵守下列一般注意事项：

- 在开始操作之前，必须完全清楚所用物质的特性。
- 没有警告并不表示该物质是安全的，因为信息不可能总是完美无缺的。
- 如果接触了有毒物质，应立即与当地的安全部门联系，获取指导。
- 所有化学、生物和放射性废物的处理要使用适当的处理方法。
- 关于合适手套的特殊准则，向当地安全部门咨询。
- 使用浓酸和浓碱时必须十分小心。戴上护目镜和合适的手套。如果使用量大时，要戴上面罩。

  不要将强酸和有机溶剂混合，因为它们会发生反应。特别是硫酸和硝酸会发生剧烈的放热反应，可能引起火灾和爆炸。

  不要将强碱和卤化物溶剂混合，以免反应生成可引起爆炸的活性炭烯。
- 不要用嘴对着吸管吸取溶液。这种方法既不能保证无菌，又很危险。必须使用移液装置或吸球。
- 将卤化物溶剂和非卤化物溶剂分开存放（例如，氯仿和丙酮混合后在碱存在的条件下会引发意想不到的反应）。卤化物溶剂是氯仿、二氯甲烷、三氯三氟代乙烷、二氯乙烷之类的有机溶剂。一些非卤化物溶剂有戊烷、庚烷、乙醇、甲醇、苯、甲苯、*N*,*N*-二甲基甲酰胺（DMF）、二甲亚砜（DMSO）和乙腈等。

684

- 可见或不可见的激光辐射都会对眼睛和皮肤造成严重伤害。要采取正确的预防措施，防止暴露于直射或者反射的激光束。必须严格遵守制造商的安全指南并向当地安全部门咨询。更详细的信息见下面的注意事项。
- 闪光灯由于光线很强，对眼睛可能有害。有时也会发生爆炸。要戴上合适的护目镜并遵守制造商的指南。
- 显影液和定影液中含有有害的化学物质。使用时必须按照生产商的指南进行操作。
- 如果使用不当，电源和电泳设备会造成严重的火灾和电击事故。
- 实验室中的微波炉和高压蒸汽灭菌锅需要可靠的预防措施。操作（如熔化装在瓶子里的琼脂或 Bacto 琼脂，或者是灭菌）不当时可能发生事故。如果瓶口的旋盖没有完全松开，没有足够的空隙排出蒸汽，从微波炉或高压蒸汽灭菌锅中取出瓶子时会发生爆炸并造成严重的伤害。用微波炉加热或进行高压蒸汽灭菌时必须将瓶盖完全松开。在不需要对琼脂灭菌的常用琼脂糖凝胶中，替代方法是量取需要的琼脂并将溶液加入到烧瓶中。
- 使用切片刀、解剖刀、剃刀或注射针等切割装置时要非常小心。切片刀非常锋

利！切片时要小心操作。如果使用不熟练，应该请人示范正确的操作方式。废弃物处置时，要放在实验室中专用的“利器”处理容器内。使用过的注射针处理时不要给注射针套上针套，但是注射针仍然要装在注射器上而不要取下。这样可以避免操作时发生伤害（和可能的感染；见生物安全性），因为试图套回针套时往往会发生事故。破碎的巴斯德吸管，盖玻片或载玻片也都可能会带来伤害。

常用化学品的一般特性

危险品目录分类总结如下：

- 无机酸，如盐酸、硫酸、硝酸或磷酸，都是带有刺激性烟雾的无色液体。要避免溢出物粘到皮肤或衣物上。如果溢出，要用大量的水进行稀释。这类浓酸会毁坏纸张，织物和皮肤，以及对眼睛造成严重伤害。
- 无机碱，如氢氧化钠是白色固体，溶于水并放热。浓溶液会缓慢地溶解皮肤甚至指甲。
- 重金属盐通常是可溶于水的有色粉末状固体。大多数重金属盐是强的酶抑制剂，对人和环境（如鱼和藻类）有毒害。
- 大多数有机溶剂是可燃性的易挥发液体。要避免吸入挥发性气体，否则会引起恶心或头晕。也要避免接触皮肤。
- 其他有机化合物，包括有机硫化物如巯基乙醇或者是有机胺，气味非常难闻。其他的则是高反应活性的，使用时要十分的小心。
- 如果操作不当，染料及其溶液不仅会对样品染色，还会造成皮肤和衣物的染色。其中一些染料是诱变的（如溴化乙锭）、致癌的和有毒性的。

- 所有以“ase”为结尾的名称（如过氧化氢酶、β-葡糖醛酸糖苷酶或消解酶）都指的是酶。也有一些酶的命名不是系统命名，如胃蛋白酶。它们中的大多数由生产商提供时都含有缓冲物质等。要注意这些物质各自的特性。
- 有毒化合物常用在对细胞的处理中。它们有潜在的危险，必须正确操作。
- 注意对一些所列出化合物的毒理特性尚未完全研究清楚。使用各种化学品时都要十分谨慎。尽管化合物的毒性作用可以定量表示（如 LD_{50} 值），但是致癌剂或诱突剂却不同，单次接触就可能造成伤害。同时也要注意到一种给定化合物的危险性与其物理状态有关（细粉状 vs. 大块晶体/乙醚 vs. 甘油/干冰 vs. 装在压力容器内的二氧化碳）。预料实验中在什么情况下最可能暴露在这些物质之下，以及如何使自己和环境得到最佳的保护。

危险品

[α-^{32}P] CTP，见放射性物质

乙酸（浓）的腐蚀性很强，用时要小心。液态和气雾状乙酸都会严重烧伤全身组织。吸入、摄取或者经皮吸收都会造成伤害。戴上手套和护目镜，在化学通风橱内使用。远离热源、火花和明火。

乙酸酐对皮肤、眼睛、黏膜和上呼吸道极其有害。吸入、摄取或者经皮吸收都会造

成伤害。戴上手套和护目镜，在化学通风橱内使用。

乙腈（甲基氰）极易挥发和特别易燃。它是一种刺激物和化学窒息剂，吸入、摄取或者经皮吸收都会造成伤害。大量接触的处理办法同氰化物中毒处理一样。戴上手套和护目镜，在化学通风橱内使用。远离热源、火花和明火。

丙烯酰胺（非聚合的）是一种强烈的神经毒剂，可以通过皮肤吸收（其效应是累加性的）。避免吸入粉末。称量粉末状的丙烯酰胺和亚甲基双丙烯酰胺时要戴上合适的手套和防护面具。在化学通风橱内操作。聚丙烯酰胺通常认为是无毒性的，但是操作时也要小心，因为其中可能会含有少量非聚合的丙烯酰胺。

S-腺苷甲硫氨酸（SAM）具有毒性，吸入、摄取或者经皮吸收都会造成伤害。戴上手套和护目镜，在化学通风橱内使用。避免吸入粉末。

乙酸铵（$H_3CCOONH_4$），吸入、摄取或者经皮吸收都会造成伤害。戴上手套和护目镜，在化学通风橱内使用。避免吸入粉末。

过硫酸铵［$(NH_4)_2S_2O_8$］，对黏膜和上呼吸道组织、眼睛和皮肤有极大的危害性。吸入可致命。操作时戴上手套、护目镜和防护服，只能在化学通风橱内使用。操作后彻底洗净。

溴酚蓝，吸入、摄取或者经皮吸收都会造成伤害。戴上手套和护目镜，在化学通风橱内使用。

含砷的二甲胂酸盐是强毒剂，吸入、摄取或者经皮吸收都有可能致命。戴上手套和护目镜，在化学通风橱内使用。

氯仿（$CHCl_3$）对皮肤、眼睛、黏膜和呼吸道有刺激作用。它是一种致癌剂，对肝和肾有损伤。它也是挥发性的，应避免吸入蒸气。戴上手套和护目镜，在化学通风橱内使用。

柠檬酸是一种刺激物，吸入、摄取或者经皮吸收都会造成伤害，它有对眼睛造成严重伤害的危险。戴上手套和护目镜。避免吸入粉末。

环乙酰亚胺吸入、摄取或者经皮吸收都有可能致命。戴上手套和护目镜，在化学通风橱内使用。

DAB，见 3,3′-二氨基联苯胺四盐酸

DEPC，见焦碳酸乙二酯

3,3′-二氨基联苯胺四盐酸（DAB）是一种致癌物。使用时要非常小心。避免吸入蒸汽。戴上手套和护目镜，在化学通风橱内使用。

焦碳酸乙二酯（DEPC）是一种强力的蛋白质变性剂，并且被怀疑致癌。开瓶时使瓶子远离操作者；瓶内压力会使液体飞溅。戴上手套和护目镜，穿上实验服，在化学通风橱内使用。

***N*,*N*-二甲基甲酰胺**（DMF）［$HCON(CH_3)_2$］，可能有致癌作用，对眼睛、皮肤和黏膜均有刺激性。通过吸入、摄取或者经皮吸收都发挥其毒性作用。长期吸入会导致肝、肾损伤。戴上手套和护目镜，在化学通风橱内使用。

二甲基亚砜（DMSO）可因吸入或经皮吸收而产生危害。戴上手套和护目镜，在化学通风橱内使用。DMSO 还是易燃的。储存在密封容器内。远离热源、火花和明火。

二硫苏糖醇（DTT）是一种强还原剂，散发有难闻气味。吸入、摄取或者经皮吸收都会造成伤害。使用固体或者高浓度储存液时，要戴上手套和护目镜，在化学通风橱内使用。

DMF，见 *N*,*N*-二甲基甲酰胺

DMSO，见二甲基亚砜

DTT，见二硫苏糖醇

溴化乙啶是一种强烈的诱变剂，并且有毒。向当地研究机构的安全官员咨询专门的使用和废弃方法。避免吸入粉末。使用含本染料的溶液时必须戴上手套。

甲醛（HCHO），毒性大且易挥发。也是一种致癌剂。它很容易通过皮肤吸收，对皮肤、眼睛、黏膜和上呼吸道有刺激性或损伤。避免吸入蒸汽。戴上手套和护目镜，在化学通风橱内使用。远离热源、火花和明火。

甲酰胺是一种致畸剂，其蒸气对眼睛、皮肤、黏膜和上呼吸道有刺激性。吸入、摄取或者经皮吸收都会造成伤害。使用高浓度甲酰胺溶液时必须戴上手套和护目镜，在化学通风橱内使用。工作溶液尽量要盖好。

HCl，见盐酸

HCHO，见甲醛

H_3COH，见甲醇

H_2O_2，见过氧化氢

$HOCH_2CH_2SH$，见 β-巯基乙醇

盐酸（HCl），易挥发，吸入、摄取或者经皮吸收都可能致命。对黏膜、上呼吸道、眼睛和皮肤有极强的破坏性。戴上手套和安全眼镜，在化学通风橱内小心使用。使用量大时必须戴上护目镜。

过氧化氢（H_2O_2），有腐蚀性和毒性，对皮肤损伤性极强。吸入、摄取或者经皮吸收都会造成伤害。戴上手套和护目镜，在化学通风橱内使用。

异戊醇易燃并有刺激性。吸入、摄取或者经皮吸收都会造成伤害。戴上手套和护目镜。避免吸入蒸汽。远离热源、火花和明火。

KOH，见氢氧化钾

激光辐射，根据所使用激光的类别，可见和不可见激光都会对眼睛和皮肤产生严重伤害，并且有可能产生空气污染。高功率激光可造成永久性眼睛损伤，烧伤暴露的皮肤，点燃易燃的物质，还能够激发有毒化学品释放有害性副产品。避免眼睛或皮肤暴露于直射或散射激光。不要用眼看激光，也不要将激光对准他人。戴上合适的护目镜，并使用针对特定的波长、工作模式（连续或脉冲式）和输出功率（瓦）激光而设计的合适防光板。不要佩戴首饰或者其他可能反射或散射激光的物体。一些非光束伤害包括电击致死、火灾和窒息。激光工作场所的入口处要严格管制并张贴警示标识表明激光正在使用。必须始终遵守仪器附带的安全指南，并向当地的安全部门咨询进一步的信息。

- 离子激光器由于使用高电压和电流的电源而容易产生危险。必须始终遵守生产商建议的安全指南。
- 紫外激光器由于光束不可见和高能量辐射而容易产生危险。必须使用消光器，

散射光防护板和荧光的光束察看卡片。

- 蓝–绿激光器由于光热致凝结作用而容易产生危险。蓝色和绿色波长的光很容易被血红蛋白吸收。

688

液态氮由于温度极低而容易导致严重伤害。操作冰冻亚样品时要非常小心。避免吸入蒸汽。液氮渗入冻存管后，在从液氮中取出冻存管时容易造成试管爆裂。可能的话，使用带O形圈的试管。戴上抗低温手套和面罩。

马来酸有毒，吸入、摄取或者经皮吸收都会造成伤害。与水或潮湿空气发生反应后会释放有毒，腐蚀性或可燃性气体。避免吸入蒸汽或粉末。戴上手套和护目镜。

MeOH 或 H_3COH，见甲醇

β-巯基乙醇（2-巯基乙醇）（$HOCH_2CH_2SH$），吸入或经皮吸收时可能致命，如果摄入也会造成伤害。高浓度β-巯基乙醇对黏膜、上呼吸道、皮肤和眼睛有强损伤性。β-巯基乙醇有非常难闻的味道。戴上手套和护目镜，在化学通风橱内使用。

甲醇（MeOH 或 H_3COH）有毒，会引起失明。吸入、摄取或者经皮吸收都会造成伤害。必须保持良好的通风，以减少暴露于蒸汽的机会。避免吸入蒸汽。戴上手套和护目镜，在化学通风橱内使用。

***N*,*N'*-亚甲基双丙烯酰胺**有毒而且会影响中枢神经系统。吸入、摄取或者经皮吸收都会造成伤害。戴上手套和护目镜。避免吸入粉末。

***N*-甲基-2-吡咯烷酮**（2-甲基-2-吡咯烷酮）吸入、摄取或者经皮吸收都会造成伤害。戴上手套和护目镜，在化学通风橱内使用。避免吸入蒸汽。远离热源、火花和明火。

3-(*N*-吗啉代)-丙磺酸（MOPS）吸入、摄取或者经皮吸收都会造成伤害。MOPS对黏膜和上呼吸道有刺激性。戴上手套和护目镜，在化学通风橱内使用。

MOPS，见3-(*N*-吗啉代)-丙磺酸

$NaBH_4$，见硼氢化钠

NaOH，见氢氧化钠

PEG，见聚乙二醇

酚具有很强的毒性和高度的腐蚀性，并会造成严重的灼伤。吸入、摄取或者经皮吸收都会造成伤害。戴上手套、护目镜和防护服，在化学通风橱内使用。接触到酚的任何皮肤部位要用大量的水冲洗，并用肥皂和水清洗。切勿使用乙醇！

苯甲基磺酰氟（PMSF）（$C_7H_7FO_2S$ 或 $C_6H_5CH_2SO_2F$），是一种剧毒的胆碱酯酶抑制剂。它对呼吸道黏膜、眼睛和皮肤都有极强的破坏性。吸入、摄取或者经皮吸收都可能致命。戴上手套、护目镜，在化学通风橱内使用。一旦接触，要立即用大量的水冲洗眼睛或皮肤，并丢弃被污染的衣物。

PMSF，见苯甲基磺酰氟

689

聚乙二醇（PEG）吸入，摄取或者经皮吸收都会造成伤害。避免吸入粉末。戴上手套和护目镜。

二甲胂酸钾，见二甲胂酸盐

氢氧化钾，KOH 和 KOH/甲醇，有剧毒，吞服会致命。吸入、摄取或者经皮吸收都会造成伤害。溶液有腐蚀性，会引起严重烧伤。操作时要非常小心。戴上手套

和护目镜。

放射性物质：在计划需要使用放射性物质的实验时，应该考虑同位素的理化特性（半衰期、射线类型和能量）、放射性物质的化学形态、放射浓度（比活）、总量及其浓度。只订购和使用所需要的用量。操作放射性材料时必须始终戴上手套和护目镜，穿上实验服。专用装置或放射性物质产生的 X 射线和 γ 射线都是波长很短的电磁波。它们从放射源各向同性地向周围发射，或者是聚焦成光束。它们潜在的危险取决于暴露时间，遭受的强度和使用的波长。要注意铅或其他类似材料才是合适的防护。防护层的厚度视 X 射线和 γ 射线的强度而决定。有关正确使用和处理放射性废物更为详细的信息，请向当地安全部门咨询。放射性同位素使用后必须彻底检测和监控。有关常规放射性计算的便利计算器见下面的网址：http://www.graphpad.com/calculators/radcalc.cfm

SAM，见 S-腺苷甲硫氨酸

SDS，见十二烷基硫酸钠

硼氢化钠（$NaBH_4$）有腐蚀性，并会引起灼伤。吸入、摄取或者经皮吸收都会造成伤害。戴上手套和护目镜，在化学通风橱内使用。

十二烷基硫酸钠（SDS）有毒，是一种刺激物，对眼睛有严重伤害。吸入、摄取或者经皮吸收都会造成伤害。戴上手套和护目镜。不要吸入粉末。

氢氧化钠（NaOH），含有 NaOH 的溶液都有很强毒性和苛性，操作时必须非常小心。戴上手套和面罩。其他浓碱也应该按照同样的方法使用。

琥珀酸酐是一种可能的诱变剂，对眼睛有强刺激性。吸入、摄取或者经皮吸收都会造成伤害。戴上手套和护目镜，在化学通风橱内使用。不要吸入粉末。

TCA，见三氯乙酸

三氯乙酸（TCA）有很强的腐蚀性。戴上手套和护目镜。

胰蛋白酶会引起过敏性呼吸道反应。吸入、摄取或者经皮吸收都会造成伤害。不要吸入粉末。戴上手套和护目镜。使用时保证通风良好。

690

紫外光和（或）紫外辐射有危害性，会损伤视网膜。切勿用裸眼直视未经防护的紫外光源。实验室中常用的紫外光源包括手持紫外灯和紫外透射仪。必须通过可吸收有害波长的滤光片或护目镜进行观察。紫外辐射也具有致突性和致癌性。为了使暴露降至最低程度，必须保证紫外光源采取了足够的防护。在紫外光源下操作时必须戴上合适的防护手套。

二甲苯是易燃性的，高浓度时有麻醉作用。吸入、摄取或者经皮吸收都会造成伤害。戴上手套和护目镜，在化学通风橱内使用。远离热源、火花和明火。

二甲苯苯胺，见二甲苯

附录4 供　应　商

除了文中列出的那些供应商和地址外，本书中提及的所有供应商都可以在 *BioSupplyNet Source Book* 中找到，也可以查寻它们的网址：

http://www.biosupplynet.com

可通过下列方法之一免费获取 *BioSupplyNet Source Book* 目录：

- 在网站 http://www.biosupplynet.com 填写免费目录索取表格
- 发送电子邮件至 info@biosupplynet.com 索取
- 传真至 1-919-659-2199 索取

DNA 微阵列网站*

DNA Microarrays：*a molecular cloning manual* 一书的网站（www. dna-microarrays. org）提供了关于这个发展迅速的研究领域的附加信息。该站点包括：

- 印刷版中由 Piero Carninci 编写的附录 1：构建全长基因文库的扩充材料。
- 链接到 Medline 的参考文献。
- 对一线科学家有价值的其他数据库链接。
- 用于疑难解答且选自书中的图片。

补充信息将会在本书出版后添加。要访问该站点，请按如下步骤进行：

1. 打开站点主页。
2. 进行简单注册。由于网站对所有完成注册手续的个人开放，所以不需要特别的访问口令。
3. 你的 E-mail 地址和密码（注册过程中选定）即为以后访问该网站的登录信息。
   网站的 FAQ 部分包含了关于注册手续的回答。
   关于注册的其他帮助以及其他所有关于 dna-microarrays. org 网站的咨询，请发邮件给 support@dna-microarrays. org，或者在美国东部时间上午 8：00 到下午 5：00 致电 1-800-843-4388（在美国本土和加拿大）或 516-422-4100（其他所有地区）。

* 英文原著配有相应网站，为读者提供相关领域内的信息。——出版者

中文版索引

（按汉语拼音排序）

原 著 索 引

（加中文注释）

* 此索引中标注页码，均为原书页码。请读者参见正文版心外标注的页码。——出版者

[General Information]
书名=DNA微阵列实验指南
作者=（美）D．鲍特尔，（美）J．萨姆布鲁克著
页数=691
SS号=14076066
DX号=
出版日期=2016.06
出版社=北京科学出版社